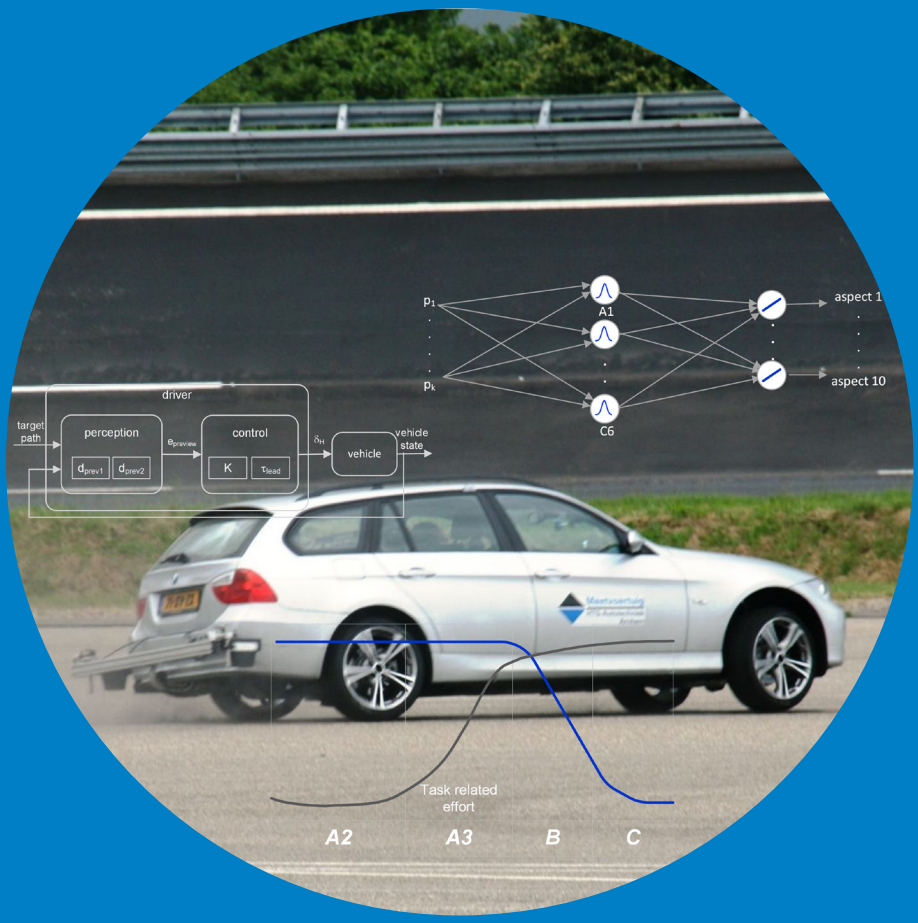
Department of Engineering Design and Production

Feel the Tire

Tire Influence on Driver's Handling Assessment

Saskia Monsma



Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 202/2015

Feel the Tire

Tire Influence on Driver's Handling Assessment

Saskia Monsma

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Engineering, at a public examination held at the lecture hall T2 of the Computer Science Building, Konemiehentie 2, 02150, Espoo, on 27 November 2015 at 12:00 h.

Aalto University
School of Engineering
Department of Engineering Design and Production
Automotive Engineering

Supervising professors
Professor Matti Juhala, Aalto University, Finland
Professor Kari Tammi, Aalto University, Finland

Thesis advisor
Professor Joop Pauwelussen, HAN University of Applied Sciences, The Netherlands

Preliminary examiners
Professor Annika Stensson Trigell, KTH Royal Institute of Technology, Sweden
Professor Karel Brookhuis, University of Groningen, The Netherlands

Opponents
Professor Annika Stensson Trigell, KTH Royal Institute of Technology, Sweden
Dr. Otto Lappi, Helsinki University, Finland

Aalto University publication series
DOCTORAL DISSERTATIONS 202/2015

ISBN 978-952-60-6547-2 (printed)
ISBN 978-952-60-6548-9 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-6548-9

Images: Front picture: Xera A. Alberts MA

Unigrafia Oy
Helsinki 2015

Finland

Publication orders (printed book):
s.monsma@han.nl

441 697
Printed matter

Aalto University

Abstract

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Author
Saskia Monsma

Name of the doctoral dissertation
Feel the Tire
Tire Influence on Driver's Handling Assessment

Publisher School of Engineering

Unit Department of Engineering Design and Production

Series Aalto University publication series DOCTORAL DISSERTATIONS 202/2015

Field of research Automotive Engineering

Manuscript submitted 2 September 2015 **Date of the defence** 27 November 2015

Permission to publish granted (date) 4 November 2015 **Language** English

☒ **Monograph** ☐ **Article dissertation (summary + original articles)**

Abstract

A key question in the development of a tire is "How can this tire improve vehicle handling?" Good handling tires contribute not only to active safety of vehicles, but also to the pleasure of driving. Handling performance is largely determined by the driver. Therefore, the final and most important handling assessment of tires is done by professional test drivers driving on a handling circuit and giving their subjective opinion. This provides the tire manufacturer with the important tire handling performance, but it gives limited information on *what* the driver perceives as good and *how* his opinion is formed; the driver is still a 'black box'.

Three methods, all based on field experiments, for gaining this knowledge about subjective assessments were chosen for this research. They have in common that they predict the driver's subjective assessment of tire handling, based on vehicle dynamics measurements. The differences lie in the way they derive and utilize these measurements. For method 1, the prediction is done with a General Regression Neural Network based on vehicle dynamics measurements. With this method, several limitations for using regression for tire handling can be circumvented.Method 2 focuses on the driver's workload as an indication for his subjective assessment. This method derives from the fact that the driver adapts to changing vehicle handling behavior. Method 3 also focuses on the driver but not by looking at measures from 'outside' the driver, like workload measures, but by modeling the driver behavior during closed-loop driver-vehicle simulations and looking at driver parameters 'inside' the driver (model).

The results show that all three methods can predict the driver's opinion about tire handling, based on vehicle dynamics measures. Analysis of the relevant measures for the prediction of methods 1 and 2, provides information on the 'what'-question. Likewise, method 3 provides information on the 'how'-question. In addition, drivers adapting behavior, e.g., compensating for less good handling tires by investing more effort, can be quantified with the mental workload measures. This makes them good indicators of driver's perceived tire handling behavior, even when the performance measures do not show differences. For implementation of one or more methods, only a subset of the vehicle dynamics measurements used for this research is needed. During use, the methods can be adapted to changing tire testing methods, with different measurements, handling aspects or maneuvers. When vehicle and tire models are available, these methods can also be used for virtual testing, predicting driver's opinion on tire handling. This research provides a first step in opening up the 'black box' of the driver by quantifying the driver's tire feeling.

Keywords tire characteristics, tire, handling, assessment, driver modelling, driver mental workload, general regression neural network, vehicle, dynamics

ISBN (printed) 978-952-60-6547-2 **ISBN (pdf)** 978-952-60-6548-9

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2015

Pages 272 **urn** http://urn.fi/URN:ISBN:978-952-60-6548-9

Acknowledgements

This research was carried out in the HAN Automotive Mobility Research group of the HAN University of Applied Sciences in cooperation with the Vehicle Engineering Research Group of the Department of Engineering Design and Production at Aalto University and Apollo Vredestein B.V., for which all three are gratefully acknowledged.

I want to thank Paul Bremmer and Gerlof Korte from Appollo Vredestein, for providing me with this opportunity to do this research, their support, knowledge and, last but not least, many, many tires. In addition, the professional tire test drivers Gerie Oude Tanke and Tjeerd Buijinck I would like to thank, who not only provided excellent work, but also proved to be very skilled in repairing vehicle suspensions and measurement equipment. They also showed my students, colleagues and me how to properly change tires very fast. Also, thank you, Gerie, for making me very carsick on the Nürburgring.

I want to thank Noldus for providing me with measurement equipment and software for the driver measurements and Tobias Heffelaar for being available for support and discussion.

I am deeply indebted to Professor Matti Juhala for giving me the opportunity to do this doctoral thesis at Aalto University, for working there, for his support and encouragement and the beautiful sightseeing tours at Helsinki and the surroundings he provided me with. He more than once reminded me that there is a whole life of research *after* a PhD. I am grateful for the support I received from Professor Kari Tammi, for his valuable guidance and feedback.

Professor Joop Pauwelussen, I cannot thank him enough; he has been the initiator of this research, he recruited me, he motivated me, he pushed me when I needed it, he protected me, mainly from myself; but most of all, he let me do my research. So often he took over my 'other job', lectures, projects, *etc.* so I could focus on my research. Discussion with him was always in-depth, positive critical and sharpened my focus. I can state that without him, I could not have done this.

I also want to thank 'my' students, students performing their research project or master thesis project in this research. First, there was Laura Luijten, such a clever and lively student that was passionate about cars. She really got me started with her difficult questions about vehicle dynamics. Lieke Arts started with no automotive background and amazed everyone by gathering so much knowledge and simulation skills in only a few months' time. I am still jealous at

her project management skills. Daan Heijne measured everything there is to measure on our test vehicle and assisted on the test track. Shrey Sultania told me a lot about India and could point out the differences with The Netherlands with much humor. Enrico Defilippi worked with me on the GRNN and his work was later take over by Venkat Kovvuru. Both students could combine their automotive background very well with to them both new area of neural networks. Martijn van Oort was my student for the longest time, next to his research project, he also did his Master thesis project in this research. He was very dedicated, smart and could really dig into things. Therefore, he was also a good discussion partner. I have learned a lot from his programming skills.

Many thanks to everyone preparing and assisting the field testing; these days went smooth, were filled with hard work, but above all were very motivating.

My colleagues are much appreciated, not only because of their feedback, discussion and sharing of knowledge, but also for motivating me.

Finally, I want to thank my friends and family for being so supportive and patience all these years, for allowing me to be absent so many times on social occasions, so I could work on my research. Thank you.

Contents

Lists of Symbols and Abbreviations

Latin Symbols

$\boldsymbol{a}$	Activation vector
a_x	Vehicle longitudinal acceleration in CoG
a_y	Vehicle lateral acceleration in CoG
$A2$	Mental workload region: high performance, low workload region
$A3$	Mental workload region: task related effort region
A_{sine}	Sine amplitude
b	Bias value
B	Magic formula tire model stiffness factor coefficient
B	Mental workload region: overload region
$\boldsymbol{c}$	Center point (one or multidimensional)
C	Magic formula tire model shape factor coefficient
C	Mental workload region: low workload region
C_α	Tire cornering stiffness
D	Magic formula tire model peak factor coefficient
d_{prev1}	Preview distance 1
d_{prev2}	Preview distance 2
$d_{preview}$	Preview distance
e_{actual}	Actual path error
$e_{preview}$	Preview path error
E	Magic formula tire model curvature factor coefficient

f	(Non)linear regression function
F	Ground reaction tire force vector
F_x	Longitudinal tire force vector
F_y	Lateral tire force vector
F_z	Vertical tire force vector
$F(\boldsymbol{p})$	Error function of driver model parameter vector $\boldsymbol{p}$
$G(s)$	Laplace transform of dynamic system behavior
H_0	Steady state gain
$H(s)$	Laplace transform of human describing function
$\boldsymbol{iw}$	Input weights vector
K	Driver model proportional gain
M	Ground reaction tire moment vector
M_H	Steering wheel moment
M_x	Overturning tire moment
M_y	Tire moment, including rolling resistance, braking or driving moment
M_z	Aligning tire moment
$\boldsymbol{ow}$	Output weights vector
$\boldsymbol{p}$	Vector of predictors
$\boldsymbol{p_A}, \boldsymbol{p_B}, \boldsymbol{p_C}$	Parameter vector for resp. driver model A, B or C
r	Yaw velocity; angular velocity around z-axis
r	Euclidian distance
r	Pearson's correlation coefficient
R^2	R-squared criterion
s	Laplace variable
s	Sample standard deviation
sp	Spread value of radial basis neuron
$\boldsymbol{t}$	Target vector
$\bar{t}_n$	Mean model prediction

t_p	Pneumatic trail tire
T_{eq}	Equivalent time, the time derived from the frequency where phase lag equals 45 degrees.
T_{sine}	Sine period
v	Horizontal velocity
v_x	Longitudinal velocity
v_y	Lateral velocity
(x, y)	Local vehicle fixed axis system with origin at vehicle CoG
x	Vehicle longitudinal axis
$\boldsymbol{x_i}$	Vector of predictor variables for observation i
(X, Y)	Global, earth fixed axis system
X_{prev1}	Position in 1st lane when driver looks at 2nd lane
X_{prev2}	Position in 2nd lane when driver looks at 1st lane
$(\boldsymbol{x}_i, \boldsymbol{t}_i)$	Data point, combination of predictors and targets for observation i.
y	Vehicle lateral axis
$\boldsymbol{y}$	Output vector
$Y_{model,n}$	n^{th} sample of the vehicle model Y position
$Y_{test,n}$	n^{th} sample of the measured vehicle Y position
w_i	ANN ith connection weight
z	Vehicle vertical axis

Greek Symbols

α	Tire slip angle
β	Vehicle side slip angle in CoG
γ	Tire camber angle
δ_H	Steering wheel angle
$\delta_{H,model,n}$	n^{th} sample of the δ_H driver model output
$\delta_{H,test,n}$	n^{th} sample of the δ_H measured output

ε_i	Error between the calculated model outcome variable and the target variable of observation i
θ	Scalar or vector of unknown parameters in regression model
μ	Normalized lateral force
μ_p	Peak value of the normalized lateral force μ
σ	Tire relaxation length
$\sigma_{\delta_H,test}$	Standard deviation of the measured steering wheel angle
$\sigma_{Y,test}$	Standard deviation of the measured Y position of the vehicle
τ	Human reaction time delay
τ_e	Total delay time constant
τ_N	Human neuromuscular lag time delay
τ_{lag}	Lag time constant
τ_{lead}	Lead time constant
$\tau_{preview}$	Preview time
τ_{prev1}	Preview time 1
τ_{prev2}	Preview time 2
φ	Neuron activation function
ψ	Vehicle yaw angle
ω_c	Crossover frequency
Ω	Wheel rotation velocity

Abbreviations

A1	Assessment tire 1 of batch A
A2	Assessment tire 2 of batch A
ANN	Artificial Neural Network
B2	Assessment tire 2 of batch B
B3	Assessment tire 3 of batch B
B4	Assessment tire 4 of batch B

C5	Assessment tire 5 of batch C
C6	Assessment tire 6 of batch C
CoG	Center of Gravity
ECG	Electro cardio graph
GRNN	General Regression Neural Network
HF	High Frequency
HFA	High Frequency Area
HRV	Heart Rate Variability
IBI	Inter Beat Interval
LF	Low Frequency
PSD	Power Spectral Density
RMSE	Root Mean Squared Errors
RSME	Rating Scale Mental Effort
SEM	Standard Error of the Mean
SSE	Sum Squared Errors
SST	Sum Squared Total

List of Figures

Author's Contribution

The author has done all the research described in this thesis herself, with the exception of the following work done by - at that time - Master students as part of their research or Master project under the authors guidance:

- Part of the research described in chapter 8 was done by Martijn van Oort (Delft University of Technology, Netherlands). He developed the vehicle and driver models, implemented measurement equipment, designed and managed the double lane change workload experiment (described in section 4.7.5) and preprocessed the resulting data.
- Preprocessing of the data of the subjective tire handling experiment (described in section 4.6) was also done by Martijn van Oort.
- Preparing the slalom experiment (described in section 4.7.4) and performing the data preprocessing of the heart rate measurement data was done by Shrey Sultania (IIT Guwahati, India).
- Deriving the metrics for the GRNN (described in section 6.1.1) was done by Enrico Defilippi (Polytechnic of Turin, Italy).
- The driver model survey and selection described as part of section 8.1 was done by Lieke Arts (Delft University of Technology, Netherlands).

For preparation and support of the field experiments, described in chapter 4, much help was received from staff and students of the HAN University of Applied Sciences. This included preparing and supporting the experiments on the test track, instrumenting the test vehicle, driving in all the tires and assisting on the test track.

1. Introduction

1.1 Motivation

A key question in the development of a tire is “How can this tire improve vehicle handling?”. Vehicle handling can be loosely defined (Dixon, 1996) as the vehicle's 'cornering feel' and good handling improves the vehicle's active safety (Heissing & Ersoy, 2011). Next to being functional, handling is a significant part of the ‘fun factor’ of driving a vehicle (Ooki, Sakai, Muragishi, Fukui, & Ono, 2008). Good handling is therefore important; but achieving it, is no easy task. It depends on the vehicle, especially on the tires, and on the driver. These three factors are explained next

Being the only contact between the vehicle and the road, tires determine largely the handling behavior. They transfer all the forces necessary for the vehicle movement and they assist the driver in predicting the vehicle handling performance through the steering system. A tire consists of several parts (Fig. 1) that have different components, compositions and properties, which can provide the tire with certain desired characteristics, *e.g.*, low rolling resistance, good wet grip, low wear or good handling (Genta & Morello, 2009). Often, desired tire characteristics conflict with each other and consequently tire design is about finding the best compromise. What this best compromise might be depends on the specific requirements for that tire. Hence, different tires are developed for various vehicles, *e.g.,* passenger cars, motorcycles or heavy vehicles (trucks, buses). In addition, they are developed for different use, resulting in summer tires, winter tires, off-road tires, *etc.* All these tires should nevertheless contribute to good handling, often stated as ‘having good tire handling behavior’.

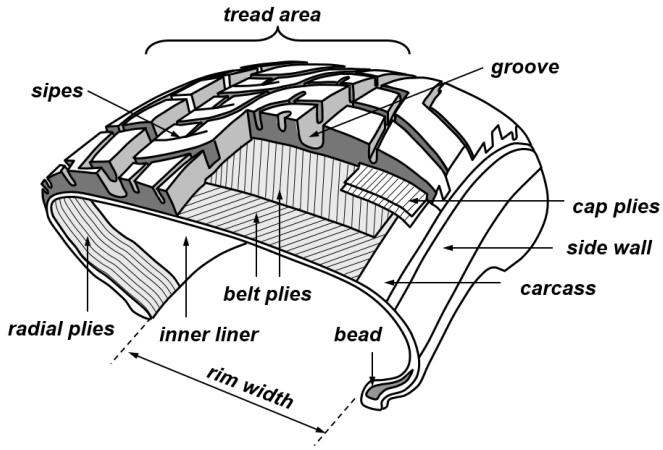


Figure 1. Exploded view of a tire (Pauwelussen, 2014).

What defines 'good handling' is not only determined by the tires, it also depends on vehicle characteristics like size, weight distribution, engine power, drive/traction configuration, suspension system. These characteristics related to the vehicle's intended use, determine the level of handling performance. For example, good handling for a sports car will be different from good handling for a saloon car.

Next to the vehicle and its tires, handling is determined by the driver, his[1] skills, preferences and expectations. This makes handling subjective. Good handling according to a rally driver will be different from good handling according to a novice driver. Handling is also influenced by the driver himself, by how he perceives the vehicle behavior and applies his controls, like steering, braking and accelerating. For example, when comparing two sets of tires, a driver may well be capable of driving the same maneuver in the same way with both sets, suggesting the tires have similar handling behavior. However, when for one of the tires sets, the driver has to apply more skills and needs more concentration to accomplish this, this can result in the tires being qualified as less good handling tires. In fact, the driver is compensating for less good tire handling behavior by putting in more effort, raising his workload.

For qualifying handling as 'good', there exist certain preferred boundary values for vehicle dynamic variables related to handling, such as reaction times, delays and gains (Pauwelussen, 2014) depending on the vehicle type, like a passenger car. By driving predefined maneuvers with different tires on a test vehicle with equipment for measuring the related vehicle dynamics (Fig. 2), one can evaluate the tire handling behavior by comparing measured or derived

[1] Drivers in general are referred to as 'the driver' and for readability, only pronouns 'he' and 'his' are used to refer to this driver, but these refer to both male and female drivers.

values with each other or with such preferred values. These measurement-based assessments are referred to as 'objective' assessments. Although these objective assessments can give a global indication of the handling behavior, this is often not enough.

As explained before, handling also includes the driver, and skilled drivers are able to distinguish small differences in handling behavior, which are not easily found with objective assessments, for example due to different tire properties. Therefore, the driver's opinion or 'subjective assessment' of handling aspects, like responsiveness, steering feel, predictability, is essential for qualifying the full handling behavior. For this reason, tire and vehicle manufacturers use subjective assessments of professional test drivers as important final handling assessments. In addition, consumer magazines, like car magazines, often have their own test drivers performing subjective (tire) handling assessments as basis for their articles. These test drivers, not necessarily professional, are highly skilled drivers able to sense and describe tire handling differences.

Figure 2. Test vehicle used for this research on a test track.

Compared to objective assessments, subjective handling assessments have some drawbacks. They are expensive, time consuming and depend heavily on the skills of the driver. Therefore, these assessments are performed only at the end of the tire development cycle as a 'sign-off' test and only with highly trained, (professional) test drivers, being able to perform tests consistently and having good skills for perceiving even small handling differences. To ensure that the evaluated handling differences are only caused by different tires, all other handling influencing factors are kept as constant as possible. This is done by using the same vehicle, maneuver, test environment and conditions. Because the weather has a major impact on tire behavior, these assessments are only done during predefined weather conditions. To keep the results comparable, the tires are tested in a limited period where often one of the tire sets is tested more than once to serve as a reference tire to indicate changing conditions.

Because these subjective assessments can only be performed with a tire at the end of the development cycle, handling problems identified with these assessments will cause the tire to go back to the design phase, raising development time and costs. With a growing demand for tires (Freedonia, 2012), a tire manufacturer is faced with a growing need for tire subjective assessments, whilst limited resources are available.

Although this subjective assessment method in itself works well, by providing the tire manufacturer with *what* the driver's opinion is, they still lack information on *how* this opinion is formed. For this, a driver is still a 'black box'. What does a driver perceive, how does this influences his opinion, what aspects are important? Answers to such questions could help the tire manufacturer in understanding the driver and including this knowledge earlier in the tire development process. Although an obvious way to get those answers seems to be 'ask the driver', this is not so straightforward. While professional test drivers can add additional comments to their subjective evaluations, it is difficult to explain this 'tire feeling' and where it comes from.

Next to this 'what' and 'how' information, tire manufacturers would like to have additional objective measures derived from subjective assessments. Such objective measures could be used to support the subjective assessments and to compare different subjective assessment results. As indicated before, subjective assessment results for the same tire often differ when they are done under different circumstances, like at another time and/or another test track. This restricts tire comparison to tires tested during the same circumstances. To loosen this restriction and to detect changing circumstances, re-testing of a tire is often used. Having additional objective measures could support this testing and could give additional information for comparing different assessment results.

Summarizing, more knowledge about subjective assessments, including driver's[2] perception, is desired by tire manufacturers to improve subjective tire handling assessments. This supports the development of good handling tires, contributing to the active safety of a vehicle and to the pleasure of driving.

Three methods for gaining this knowledge about subjective assessments were chosen for this research. All these methods have in common that they predict the driver's subjective assessment of tire handling, based on vehicle dynamics measurements, as a first step to 'open up' the black box and to quantify the driver's tire feeling. The differences lie in the way they derive and utilize these measurements.

Two often-used methods for relating objective and subjective measures in general are finding correlations between them and using general regression to predict subjective from objective measures (see section 3.1 for explanation and

[2] As tire subjective assessments can only be performed by professional or nonprofessional, but highly skilled, test drivers, these are the focus of this research and such drivers participated in the experiments. The word 'driver' or 'drivers' in this thesis when referring to this research refers to these highly skilled drivers and not to drivers in general.

references). Unfortunately, results are in most cases not applicable outside the context in which they were derived. In addition, usage of general regression require several assumptions (see section 6.2), which can hardly be fulfilled for the practical application in tire testing, *e.g.*, having a priori knowledge of the underlying model structure or having a large enough data set. Therefore, in this research, some other methods were followed.

In the first method, the driver's subjective handling assessment is predicted based on vehicle dynamics measurements with a General Regression Neural Network (GRNN). By using a GRNN for this, the previously mentioned limitations for using regression can be circumvented (see section 6.2 for explanation and references).

The second method focuses on the driver, more specifically, on his workload as an indication for his subjective assessment. This method derives from the fact that the driver adapts to changing vehicle handling behavior (see section 2.4). The assumption made and validated in this research is, that the more driver adaptation is required, the more workload the driver perceives, the lower his vehicle handling assessment score will be. Because workload cannot be measured directly, indirect measures must be used. This includes both physical and mental workload measures.

The third method also focuses on the driver but not by looking at measures from 'outside' the driver, like workload measures, but by modeling the driver behavior and looking at driver parameters 'inside' the driver (model). For this, models of the vehicle including tires are run in closed-loop with a driver model, to simulate a handling maneuver. If the driver could be modelled with parameters that are related to certain driver's characteristics, like effort or workload, this could relate to the subjective assessment of tires. If different tires (tire models) result in different (fitted) driver parameters, these 'tire dependent driver parameters' could provide information about the amount of effort and/or workload a driver perceives. This, in turn, could give a prediction of the subjective evaluation of the tire handling behavior.

1.2 Research Objectives

The previous section motivates to gain more knowledge about subjective tire handling assessment by highly skilled or professional test drivers with the aim of improving tire handling assessment methods. This main goal is detailed into the following research objectives:

- To provide information on what handling behavior is considered as good by the driver, by predicting subjective tire handling assessments based on derived objective measures and by analyzing the these measures. This information can support the actual driver's subjective evaluations and provide comparative information for assessments.
- To provide information on how the driver's subjective assessment is formed by analyzing driver model parameters from vehicle handling simulations.

1.3 Scientific Contribution

The scientific contribution of this research can be condensed into the following results:

- Tire handling assessment is found to be situated in region A3 of the performance - workload model of De Waard (Waard, de, 1996) (Fig. 19), which let driver's workload measures distinguish between tire handling differences when performance measures cannot.
- The relationship between performance and driver's workload in region A3 of the performance - workload model of De Waard (Waard, de, 1996) is found to be different for tire handling (Fig. 47). For tire handling, increasing task demand gives linearly increasing mental workload, but at a certain point the mental workload increases rapidly to maximum workload, just when the performance decreases.
- The GRNN for handling assessment has good prediction performance of driver's subjective evaluation for all 10 tire handling aspects for driving on a handling track, based on objective measures taken during only the double lane change maneuver of the handling track.
- Analysis of the used predictors of the GRNN provides information on *what* vehicle dynamic behavior is relevant for each driver. For general handling behavior, regarding all aspects, this showed that metrics based on lateral velocity, steering wheel moment and vehicle slip angle were most relevant.
- HFA is found to be the best workload measure to indicate drivers' workload differences due to speed and handling performance of tires for professional drivers. For nonprofessional drivers, this can also be seen, but this depends on the handling maneuver.
- The driver model method indicates *how* adapting driver behavior due to different tire handling performance and different task demands can be explained and how this can be related to the driver's experienced mental and physical workload.
- The lead term constant of driver model B showed to be the best candidate for predicting subjective tire handling for both professional and nonprofessional drivers.

1.4 Thesis Outline

After the introduction in this chapter, the research area background is given in chapter 2. The research approach is explained in chapter 3. Several field experiments were performed for this research, for readability, these are all described in chapter 4 and are being referred to in the following chapters. For these field experiments, assumptions were made, which are validated in chapter 5. Chapters 6 to 8 explain the three methods used in this research for predicting the driver's handling evaluation: using a GRNN, based on workload measures and based on driver model parameters. Chapter 9 discusses the overall results and conclusions.

2. Research Background

This chapter presents the research background by giving an overview of the important topics. Vehicle handling is explained in section 2.1: what is it, why is it important and when is it considered 'good'? As tires contribute significantly to vehicle handling, section 2.2 elaborates on this topic. To evaluate handling, several assessment methods exist; these are described in section 2.3. The next two sections focus on the driver, as he is an important part of handling. Section 2.4 explains the driver's adaptation to vehicle behavior and how this is related to handling performance and the driver's workload.

2.1 Vehicle Handling

In this section, first the concept 'vehicle handling' as used in this research, is defined and illustrated. The importance of vehicle handling is then explained, including criteria for good handling.

Vehicle handling is a widely discussed topic, not only in technical and scientific context, but also in normal day life. It is a broad concept, with different interpretations, but there is common agreement on the fact that it has to do with the 'cornering feel' of the vehicle. The definition of vehicle handling adopted for this research is (Dixon, 1996):

> Vehicle handling is the ability of a car to round corners successfully, the study of how this occurs and the study of the driver's perception of the vehicle cornering behavior.

This definition emphasizes that vehicle handling is not only determined by the vehicle behavior, but also by the driver's perception of this behavior, the 'cornering feel'. It also incorporates vehicle handling being subjective; similar vehicle cornering behavior can be perceived differently by drivers, depending on their experience, preferences and expectations. This also relates to the effort the driver has to put in to obtain this vehicle behavior. In early research, it was stated (Bergman, 1973) that 'the driver plays by far the most important role in vehicle handling'. Vehicle handling is therefore not seen as a property of a vehicle alone, but as a property of the total (closed-loop) driver-vehicle system as depicted in Fig. 3. Based on various inputs, as the path to follow, observed road and traffic conditions but also the 'feel' of the vehicle (through steering wheel, pedals, forces felt, sound) the driver controls the vehicle through the

steering wheel, gear, brake and accelerator pedal. These control actions, together with environmental vehicle inputs, labelled as 'disturbances', like wind force or road slope, will result in certain vehicle behavior, for example the driven path, the speed, acceleration, noise, vibrations. The driver perceives this vehicle behavior, which gives him feedback on the result of his control actions, on basis of which he will adjust his control actions. This closes the loop as shown in the figure.

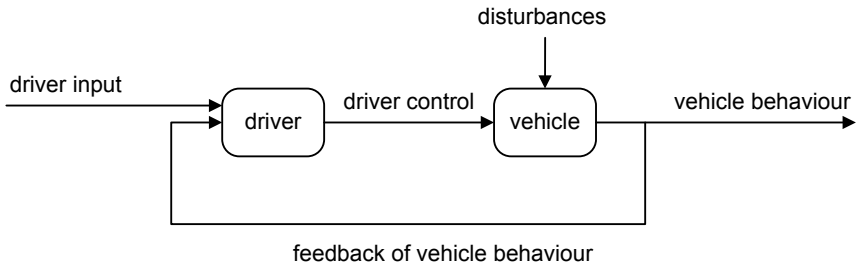


Figure 3. Closed-loop driver-vehicle system.

2.1.1 Importance of Vehicle Handling

Handling is, next to braking, a major component of a vehicle's active safety (Heissing & Ersoy, 2011). In general, stopping before or driving around the object can avoid a collision. The success of this latter option is mostly determined by the handling behavior. The introduction of automatic electronic stability control in the end of the last century has increased the active safety significantly by controlling the vehicle handling behavior in critical situations (Liebemann & Meder, 2004). Despite this, driver failures in the closed-loop driver-vehicle system play the largest role in accidents (Van Elslande & Fouquet, 2007). Improving vehicle handling should therefore not only focus on vehicle performance but also on the driver.

Next to this safety aspect of vehicle handling, there is also a 'fun factor' to be considered. In addition to transport people and goods, it can also be a pleasure. Especially for the sportive driver, handling is a key issue in the fun of driving. Most drivers never, or not often, have to rely on the safety aspect of handling, but the fun aspect can be part of everyday driving. Manufacturers of vehicles and vehicle components, like tires, appeal to both of these aspects in their information to the customers. "Designed to protect you" (Apollo Vredestein BV, 2014), "Volvo. For Live" (Volvo Cars, 2014), "Freude am Fahren" (BMW AG, 2014) and "Drive the Change" (Renault, 2014) are some typical examples. This also applies to the automotive magazines; next to the safety aspect, for example dominant in winter tire tests, focusing on the pleasure of handling is frequently done, vehicle comparison tests being a good example. In research, this aspect of handling is rarely explicitly mentioned, although there are exceptions (Ooki et al., 2008), like Ooki et al. stated:

> Vehicle transient response to steering . . . is an important facet in vehicle dynamics evaluation from a driver satisfaction and enjoyment standpoint.

Although rarely mentioned, it does not mean the fun factor is ignored: in evaluations of vehicle handling, where test drivers can give their opinion about various handling aspects, this factor is taken (implicitly) into account by rating evaluation aspects like 'ease of steering', 'responsiveness' and 'predictability'.

2.1.2 Good Vehicle Handling

After defining vehicle handling for this research and explaining its importance, the question remains, what is considered ‘good’ handling. Commonly agreed upon criteria of good handling are (Savkoor, Happel, & Horkay, 1999; Sharp, 1999):

- Having a short time delay between driver input and vehicle response. For example when given a steer input the vehicle should not take too long before it starts cornering.
- Having gains not too large or too small. Referring to steering, if the gain is too large, small steering input results in relative large directional vehicle behavior, which can be perceived by the driver as nervous vehicle behavior, where a too small gain can be perceived as sluggish behavior.
- Having a good compromise between stability and responsiveness. This is directly related to properly balancing the two previous criteria.
- Having the vehicle pointing in the direction of the movement, so no large vehicle slip angles.
- Having response immunity to external disturbances like side winds.
- Having small vehicle roll response during cornering.
- To receive accurate warning when approaching the limit during extreme vehicle behavior, so a driver can anticipate in time.
- Having consistent vehicle behavior in changing conditions, like speed, loading, and road conditions. This is in contradiction with the previous criterion; therefore, consistent behavior should be limited to non-extreme vehicle behavior with a smooth transition to extreme behavior.

Some of the above criteria are contradictory, which makes good vehicle handling a compromise. An illustration of this is the introduction of electronic stability control in vehicles. This system can intervene in the braking system and engine power if loss of control is detected, for example during evasive maneuvering. This leads to more safety and can give a driver more confidence during cornering. On the other hand, this can likewise lead to less driver awareness of approaching the limit, because the stability control prevents familiar warnings like skidding, tire squeal or abrupt vehicle movements. If, subsequently, the vehicle actually crosses the limit, this will come as a surprise for the driver, with all its consequences. It can even lead to unwanted behavioral driver adaptation, where the driver takes more risks, because he feels protected by the system.

Handling Dependencies

It might be noticed that the given overview of good handling criteria is not defined in absolute values, but in qualifications like 'small', 'short' or 'not too large or too small'. This is because precise numerical values cannot be defined. What can be defined as good handling depends on the vehicle: the qualification of good handling for a sports car will be different from good handling for a family estate car. As tires are part of a vehicle and significantly determine the behavior of a vehicle, handling is also significantly depending on the tires.

Next to this vehicle and tire dependency, handling depends on the driver. As one driver will appreciate fast vehicle response after given a small steering input, others may find this nervous behavior and would rather have a more dampened response. This relates directly to skills, expectation and perception of the driver. He will expect different handling behavior for these two vehicles and the perception of this handling behavior compared to the expectations, also based on the driver's skills, will result in the qualification of handling for that vehicle. This does not imply that handling cannot be defined in absolute terms. For various handling aspects like reaction times and gains, there are certain preferred (ranges of) values defined for different classes of vehicles (Pauwelussen, 1999a).

An example of handling dependencies is given by Weir and Di Marco (Weir & Di Marco, 1978). They defined two handling performance metrics derived from a Bode plot of yaw rate (turning velocity) over steer input:

- H_0: steady state gain
- T_{eq}: equivalent time, the time derived from the frequency where phase lag equals 45 degrees.

They suggested plots of H_0 versus T_{eq} and have indicated areas of preferred vehicle behavior for American cars around 1978. Fig. 4 shows such a plot for European vehicles being tested around 1995 (Pauwelussen, 2014). The different ranges of preferred values for the typical and expert driver are shown which originated from Weir and Di Marco.

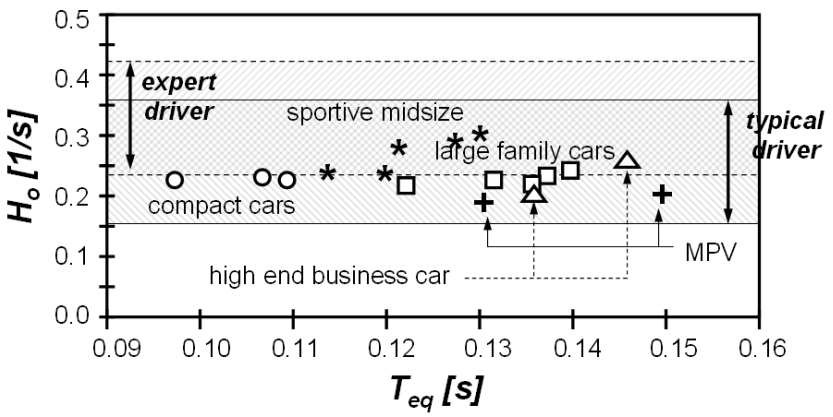


Figure 4. Weir-Dimarco plot for some European vehicles tested around 1995 (Pauwelussen, 2014).

This figure shows that the steady state gain does not vary much between vehicles. Some sportive cars having higher values, indicating that for equal steer input, the cornering radius is smaller compared to other vehicles, often referred to as ‘sportive’ steering. This figure illustrates the vehicle dependency even more by showing that preferred values for equivalent time are higher for large family cars compared to compact cars. It also illustrates driver’s dependency by showing that the preferred values for the steady state gain for expert drivers are higher than for typical drivers. In addition, it can be seen that the range of preferred values only give a global indication, within these ranges differences exist.

2.2 Tires

The significance of tires in vehicle handling is explained in this section. The term 'tire handling' is clarified in section 2.2.1 and in 2.2.2 this is done for tire design, which gives a tire its behavior or 'tire characteristics'. Tire characteristics related to handling are then described in more detail in 2.2.3. Section 2.2.4 presents linear and nonlinear tire modeling.

2.2.1 Tire Handling

With the contact between the vehicle and the outside world almost entirely depending on the tire-road contact, tires play a major role in vehicle handling. The primary forces needed for accelerating, braking and cornering are generated in the relatively small area where the tire contacts the road, the so-called tire contact patch. Therefore, tires have a considerable impact on driver's perception of vehicle handling. They affect a large part of driver's feedback of vehicle behavior (steering feel, vibrations, noise, lateral motions, *etc.*) and in predicting adverse handling conditions when the limit of controlled behavior is approached during cornering (Pauwelussen, 1999b).

The ability of the tires to contribute to vehicle handling is referred to as 'tire handling'. Tire qualifications, such as 'high performance tires’ refer to this tire handling. Obviously, a bad handling vehicle cannot become a good handling vehicle by just changing the tires, although improvement by tire adjustment is very well possible. The opposite, however, is very likely; a good handling vehicle can end up with (very) bad handling behavior by changing to worse tires (Von Glasner & Ahlgrimm, 2011), with all its consequences for ease of vehicle control and safety. Tires are therefore important vehicle components that should match the vehicle as part of the closed-loop driver-vehicle system for achieving good handling performance.

2.2.2 Tire Design

During driving, the tire contact patch is constantly subject to changes due to turning of the wheel, changing of the forces working, like tire load, and changing of the interface between tire and road, for example when driving through a

puddle. Since Dunlop developed the first practical pneumatic tire in the end of the ninetieth century (Genta, 1997), the tire has developed into a complex product consisting of more than fifty different raw materials like natural and synthetic rubber, chemicals, fiber, steel and textile. These raw materials are the basis for the various tire components, as plies, sidewalls and tread area (Fig. 5). The assembly of the various components, their use, orientation and interaction give the tire its specific characteristics (Pauwelussen, Dalhuijsen, & Merts, 2007). The relationship between tire components properties, tire characteristics and tire/vehicle handling assessment are summarized in Fig. 5. This figure also shows that certain preferred handling behavior determines requirements on tire characteristics that, in turn, determine requirements on tire components properties.

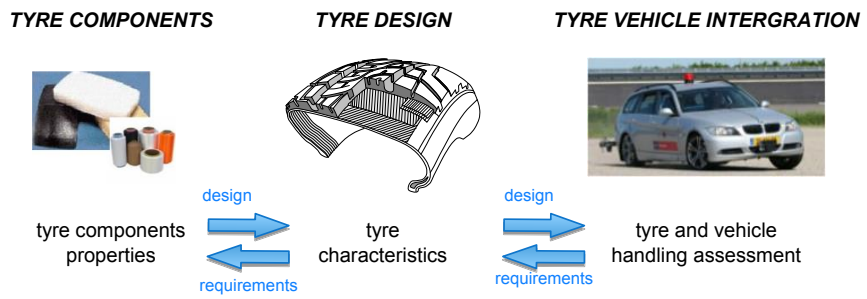


Figure 5. Relation between tire components properties, tire characteristics and tire/vehicle handling assessment.

Tire Characteristics

Tire characteristics specific for handling will be explained later, but in general, tire characteristics describe specific tire behavior, like the forces it can produce during braking, accelerating and/or cornering, its rolling resistance or the way it can dampen surface irregularities. Although called 'tire characteristics' these characteristics depend on the tire-road contact, so are also determined by the road. Most tires are designed for specific road types, like passenger car tires for paved roads and agriculture tires for unpaved surfaces. Because of this implicit assumption of the underground for a specific tire type, the actual 'tire-road' characteristics are commonly referred to as 'tire' characteristics. A tire is specifically designed to deal with changing circumstances in the tire contact patch and should be able to provide the desirable tire characteristics. What these desirable characteristics are, differ for each tire type. For a high-performance passenger car tire, we would prefer good tire characteristics related to high speed cornering, fast accelerating and braking. For winter tires, desirable tire characteristics are the ones that result in good grip during low outdoor temperatures and on snowy and icy surfaces. These are just two examples, but taken in consideration the numerous different vehicles, from motor cycles, to passenger cars and trucks, and the numerous different purposes of these

vehicles, from low-speed agricultural vehicles to race cars, it is clear that tire design is not about 'one size fits all', but is custom-made.

Even if the desirable tire characteristics can be precisely specified for a certain tire type, it is not easy to produce such a tire. Often, desirable tire characteristics do not coincide, *e.g.*, a very low rolling resistance tire is hard to combine with excellent wet grip characteristics. This results at finding the optimal compromise for each tire type. In addition, the relationship between tire components properties and its characteristics is not straightforward. Changing one component can influence several tire characteristics at the same time. These factors make tire design a complex task, where the result - the tire with its tire characteristics giving certain performance on an actual vehicle - is not entirely predictable. Therefore, testing of (prototype) tires, to assess the actual tire characteristics is an important part of the tire development. This tire testing is not done only in the laboratory with specialized tire testing equipment, but also in operating conditions with the tires mounted on a vehicle and driven on the test track as with handling tests or on the public road for certain durability tests.

2.2.3 Tire Characteristics related to Handling

Because the scope of this research is handling, in this section the main tire characteristics related to handling are described. Fig. 6 shows a schematic of a tire, using the ISO sign conventions (ISO, 1991). The tire axis system x, y and z is a right hand orthogonal system with the origin at the center of the tire road contact. The x-axis is the intersection of the wheel plane and the ground plane with a positive direction forward. The z-axis is perpendicular to the ground plane with a positive direction upwards. The y-axis is in the ground plane, positive to the left when looking in the positive x-direction, as shown in the figure. The tire is rotating with a velocity Ω. External forces acting from the ground on the tire can be summed into one ground reaction force vector F and one ground reaction moment vector M. The components in the x, y and z direction of these vectors are:

- Longitudinal tire force F_x for accelerating and braking
- Lateral tire force F_y for cornering
- Vertical tire force F_z for the load on the tire
- Overturning moment M_x
- Moment M_y , including rolling resistance, braking or driving moment
- Aligning moment M_z

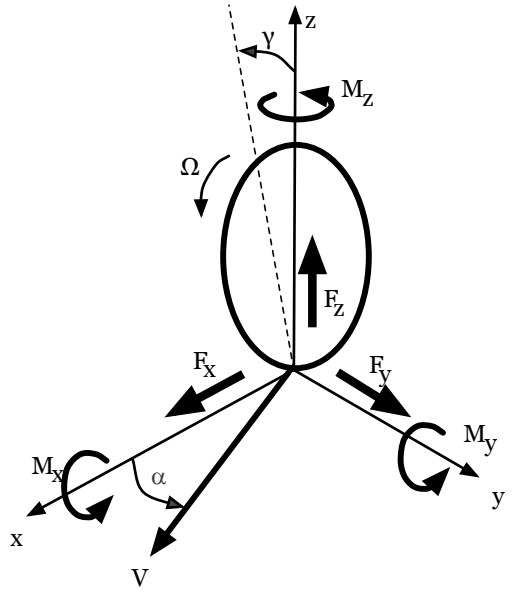


Figure 6. Forces and moments acting on the tire, tire slip angle and tire camber angle.

Two angles relevant for the tire forces and moments and shown in the figure are:

- The tire slip angle α between the velocity vector of the center of the tire contact patch and the x-axis.
- The camber angle γ between the wheel plane and the z-axis.

Steady State versus Transient Tire Behavior

Tire characteristics can be related to steady state and transient tire behavior. For the definition of steady state and transient the vocabulary of 'Road vehicles - Vehicle dynamics and road-holding ability' defined by ISO (ISO, 1991) is followed. Summarizing, steady state refers to all forces and moments in the tire road contact being constant, like driving with constant speed straight ahead or on a circle with constant radius. Transient is everything that is not steady state, like accelerating or braking, driving into a turn or coming out of a turn.

In the next parts of this section the most important steady state tire characteristics related to handling are explained first and after this the relevance of the transient tire behavior for handling is described.

Steady State Lateral Force F_y and Aligning Moment M_z

Vehicle handling is strongly determined by cornering of the vehicle. Therefore, tire characteristics important for handling are primarily related to the lateral force and aligning moment working on a tire. Fig. 7 shows a top view of a tire during cornering with slip angle α. Due to the tire elasticity, the tire will deform in lateral direction opposite to the slip angle, producing a resultant lateral force

F_y. Fig. 7 shows a tire that is cornering to the right, having a positive slip angle α (to the left) giving a (negative) lateral force to the right. These lateral tire forces will cause the vehicle to make a right turn.

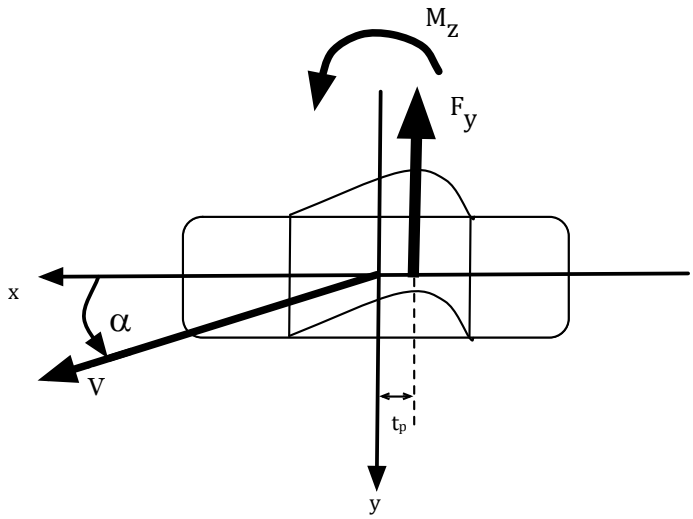


Figure 7. Top view of tire with lateral force F_y, aligning moment M_z, pneumatic trail t_p and slip angle α.

The resultant lateral force F_y is applied at a distance, the pneumatic trail t_p, behind the center of the contact zone, causing a moment:

$$M_z = t_p F_y \tag{1}$$

This is called the 'aligning' moment, because it works in the direction to align the mean plane of the wheel with the velocity vector. This moment is also felt by the driver as part of the moment felt on the steering wheel and is an important feedback signal for a driver. If the driver releases the steering wheel during cornering, the aligning moment is one of the factors that will cause the wheel to align itself with the velocity vector, causing the vehicle to start driving straight ahead. When the driver comes out of the corner, he often does not actively steer back, but automatically applies such smooth release of steering wheel to position his vehicle to straight driving.

The relationship between slip angle, lateral tire force and aligning moment is shown in two graphs of Fig. 8. The graph on the left shows the lateral force increasing almost linearly with increasing slip angle for small slip angles (until 4 or 5 degrees). For larger slip angle, the lateral force increases less than linear until reaching a peak value. After this peak, if the slip angle is further increased, the lateral force stays at its maximum or can even decrease; the latter is shown in the graph. The slope of the tangent for small slip angle is referred to as the

cornering stiffness C_α and is an important tire characteristic (Pacejka, 2012). In the graph on the right, the behavior of M_z is shown versus slip angle. For small slip angles, this is also linear, but this linear range is smaller than for F_y. After M_z reaches its maximum, it falls off with increasing slip angles to become zero or even negative. The reason for decreasing M_z, although F_z is still increasing, is the variation of the pneumatic trail with the slip angle, from a maximum value for small slip angle and becoming smaller with increasing absolute value of α. With increasing slip, the sliding part of the contact zone between tire and road expands, causing the shear stress to become more symmetric. Consequently, the resultant force F_y will then shift to the center of the tire contact patch, and t_p decreases.

Although both graphs in Fig. 8 start in the origin, giving zero lateral tire force F_y and aligning moment M_z for zero slip angle, this is a simplification. In reality, small forces and moments are often present for zero slip angle, due to irregularities and/or non-symmetry of the tire construction.

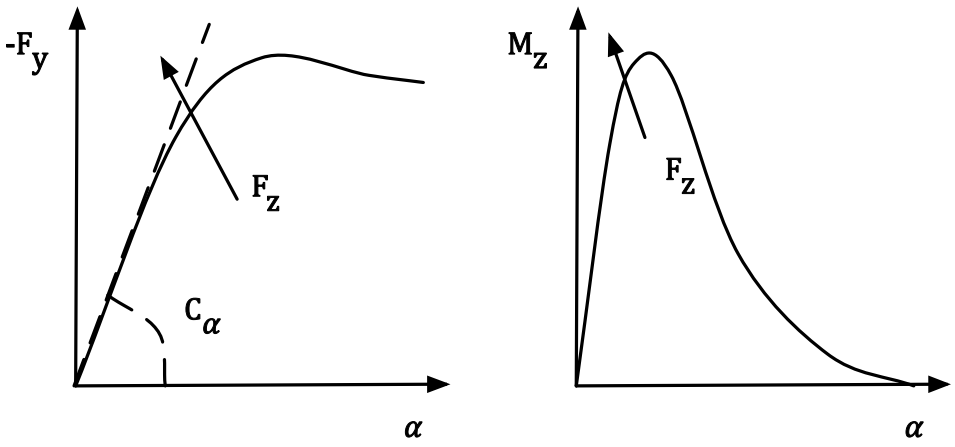


Figure 8. Lateral tire force F_y (left) and aligning moment M_z (right) as function of tire slip angle α and tire load F_z.

The lateral tire force F_y and aligning moment M_z are not only determined by the slip angle α, but also by other factors. As mentioned before, the road surface is a significant factor. Compared to a dry road, a wet, snowy, muddy or icy road can reduce the tire lateral force, especially for larger slip angles. The more substance there is between the tire and the road, the less tire-road forces can be produced.

Next to this tire-road interface, an important factor influencing vehicle steering behavior, is the tire load F_z. Fig. 8 also shows that increasing the tire load will increase the lateral tire force F_y, cornering stiffness C_α and aligning moment M_z of a tire. This increase is not linear. Fig. 9 shows the lateral tire force F_y as a function of tire load F_z for a constant tire slip angle α.

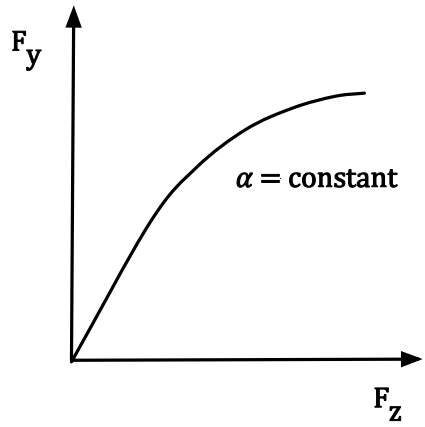


Figure 9. Lateral tire force F_y as function of tire load F_z for constant tire slip angle α.

This function is increasing less than linear, twice as much load produces less than twice as much lateral force. This has a major impact on the vehicle steering behavior during extreme cornering. During cornering, weight transfer from inner to outer wheels occurs, causing the inner wheels to be less loaded than the outer wheels compared to the static loaded situation. With the same slip angles, the less loaded inner wheels therefore produce less lateral force than the more loaded outer wheels. Due to this nonlinear behavior of tire lateral force as function of tire load, the decrease of lateral force at the inner wheels is not fully compensated by the gain of the outer wheels, resulting in a loss of total lateral force on the vehicle. This difference increases with more extreme cornering / more lateral acceleration, giving more weight transfer. Due to axles stiffness differences this loss of total lateral force can also be divided differently over the front and rear axle, affecting the steering behavior of the vehicle (more or less understeer/oversteer).

The normalized lateral force is defined by:

$$\mu = \frac{F_y}{F_z} \tag{2}$$

Another important factor influencing the tire lateral force F_y and aligning moment M_z is the presence of a longitudinal force F_x, in case of the so-called 'combined slip' situation. This situation occurs with braking or accelerating during cornering. In this situation, the lateral tire force F_y is decreased compared to the pure lateral slip situations, where the longitudinal force F_x is zero. Fig. 10 shows this influence.

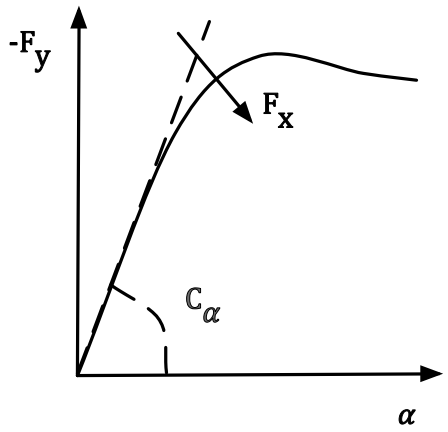


Figure 10. The influence of longitudinal force F_x on lateral tire force F_y and cornering stiffness C_α.

The aligning moment M_z is also changed in a situation of combined slip, increasing for ordinary driving force, but decreasing for extreme driving forces. For braking forces, the aligning moment M_z decreases and can even switch sign for severe braking. This means that the driver receives a moment in his steering wheel opposite to what he normally feels and expects, which can give a destabilizing effect by increasing the slip angle (Genta, 1997).

Next to the above explained three important factors influencing the lateral tire force F_y and aligning moment M_z - *i.e.*, the tire side slip angle α, the tire load F_z and longitudinal force F_x - there are other influencing factors. These can be due to the positioning of the wheel, or to irregularities in the tire design. For example, a nonzero camber angle causes an additional lateral tire force in the direction of the tilting side of the wheel. Such additional tire force may also arise when the wheels are not positioned parallel to the x-axis, like with toe-in or toe-out. Irregularities or non-symmetric tire design can give additional force and moment components to the tire, even when no slip angle is present. Compared to the aforementioned first three factors, these other factors are relatively small and are therefore often neglected.

Transient Tire Characteristic: Relaxation Length

Next to the steady state tire characteristics, dynamic tire behavior may be relevant for handling. One usually distinguishes between situations where the tire needs time to adapt to sudden changes in slip angle or external forces (F_x, F_z), and situation where the belt moves with respect to the rim. The first situation is often referred to as 'transient conditions'. If a tire is instantaneously set to a certain slip angle it takes some distance of rolling before it reaches the steady state value, see Fig. 11. This behavior is analogous to first order behavior. Similar to the time constant, the relaxation length σ can be defined as the distance a tire rolls before the lateral force F_y reaches 63% of its steady state value. At nominal vertical load the relaxation length is of the order of magnitude of the wheel radius (Pacejka, 2012).

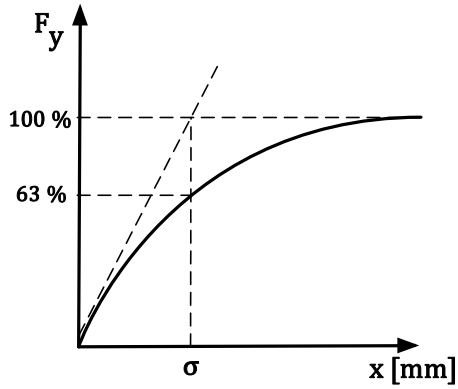


Figure 11. Relaxation length σ of a tire.

During normal driving the effect of this tire transient behavior is negligible, because the associated time delays are relatively very small during smooth steering and graduate changing conditions. During fast changing conditions, this delay can have considerable effects, especially when combined with low speed maneuvers as during parking.

2.2.4 Tire Modeling

A tire model is a mathematical representation of the behavior of a tire under various circumstances. There are many different tire models. The differences can be related to the modeling method used, going from modeling based on physical principles, resulting in so called physical or theoretical models, to modeling based on fitting to measurements, giving empirical models, with various combinations of these methods in between, often called semi-empirical models. Which tire model is most suitable, depends on the goal of the tire model usage. Physical tire models are usually simple enough to be used in mathematical analysis, and are primarily used to get more understanding of the tire behavior. They can also be used in vehicle dynamics simulations, contributing to vehicle dynamics understanding. An example is the brush model (Dugoff, Fancher, & Segel, 1970; Fiala, 1954) that implements tire deformation and resulting tire forces and moments as the bending of little beams, radially attached to a rigid carcass. Another example is the stretched string tire model (Fiala, 1954) describing primarily the carcass deflection, as a ring under tension being flexibly attached to the tire symmetry plane. Such simple physical models can give a qualitative explanation of tire behavior but are limited in giving a quantitative representation of the tire-road contact. For that purpose, more complex tire models are needed, as models based on finite element methods. Their drawback, however, is that their complexity makes them time consuming in modeling and simulation. Tire models appropriate for simulations and providing a good fit on real tire behavior belong to the (semi-) empirical models.

Examples of these models are the Magic Formula model (Bakker, Nyborg, & Pacejka, 1987; Bakker, Pacejka, & Lidner, 1989; Besselink, Schmeitz, & Pacejka, 2010; Pacejka, 2012) for steady state and transient tire behavior up to about 8 Hz, the SWIFT tire model (Besselink, Pacejka, Schmeitz, & Jansen, 2005) which model tire behavior for higher frequencies or a pure empirical model EDM (Renner & Barber, 2000).

Nonlinear Tire Model 'Magic Formula'

In this research, the semi-empirical Magic Formula model is used as tire model, because it can model linear and nonlinear tire behavior (needed for more extreme driving) very well and is also applicable for fast simulations, making it also the de facto standard tire model in research and industry for vehicle handling simulations (Besselink et al., 2010; Pacejka, 2012). It describes the longitudinal and lateral tire forces F_x and F_y and the moments M_x, M_y and M_z as functions of slip. The basic equation for the lateral force F_y as a function of the slip angle α (ignoring offsets with respect to the origin, also referred to as horizontal and vertical shifts) reads:

$$F_y = D\,sin\big(C\,arctan\big(B\alpha\ - E(B\alpha - arctan(B\alpha))\big)\big) \tag{3}$$

In the simplest form of the Magic Formula, the coefficients B, C, D and E can be taken as constant, but in the more enhanced version the coefficients depend on the tire load F_z, the tire camber angle γ and on tire pressure. In addition, combined slip may be accounted for. This dependency is defined with various parameters in the underlying expressions for these coefficients. This results in a well-defined set of parameters for the Magic Formula, describing the tire behavior in detail. The cornering stiffness can be calculated by:

$$C_\alpha = BCD \tag{4}$$

From Fig. 8 it can be seen that the force and moment characteristics are linear for small slip angle values. For normal every day driving, the slip angles remain in this linear area. For modelling this behavior, a linear tire model is often used. Using (1), (2) and (3), these linear tire models are easily found to be:

$$F_y = C_\alpha \alpha \tag{5}$$

$$M_z = t_p C_\alpha \alpha \tag{6}$$

For this research, the full nonlinear tire model is used, because during the experiments extreme driving is performed. Referring again to the nonlinear Magic Formula model, the peak factor D can be written as

$$D = \mu_p F_z \tag{7}$$

with μ_p being the peak value of the normalized lateral force μ (Eq. (2)).
The shape factor C influences the decrease of F_y beyond the maximum value, if it is monotonously increasing, if it shows a local maximum, and to what extent F_y will decay beyond this maximum.

Consequently, the so-called stiffness factor B can be derived from the other values

$$B = \frac{C_\alpha}{CD} \tag{8}$$

and it determines the cornering stiffness of the tire.

The curvature factor E determines the curve of the graph after the maximum value of F_y. The relationship between the graph of $F_y(\alpha)$ and the coefficients of the Magic Formula is shown in Fig. 12.

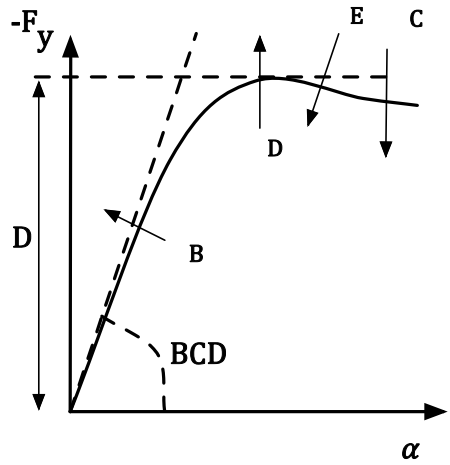


Figure 12. Magic Formula coefficients for lateral force F_y.

To tune the tire characteristics for circumstances other than during the measurements, like different road surface, friction properties, tire wear, temperature, humidity, scaling factors have been defined for this tire model (Pacejka, 2012). The most important scaling factors related to handling are those for the lateral peak friction coefficient D, the cornering stiffness C_α and the pneumatic trail t_p.

2.3 Handling Assessment

Vehicle handling assessments can be divided in different ways, which will be explained next:

- How the vehicle is controlled, open-loop versus closed-loop control.
- How the vehicle behavior is assessed, objective versus subjective.

- How tests are performed, real versus virtual testing.

2.3.1 Open-loop versus Closed-Loop Assessment

Assessment of vehicle handling can be divided in how the vehicle is controlled, open-loop or closed-loop. Open loop steering input is pre-determined. Therefore, open-loop assessment only concerns assessment of the vehicle response, shown in Fig. 13.

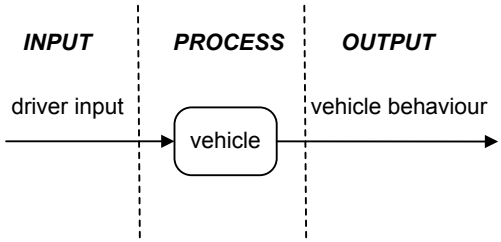


Figure 13. Open-loop assessment.

Closed loop steering input depends on vehicle response. Therefore, closed-loop assessment concerns the total closed-loop driver-vehicle system performance, shown in Fig. 14.

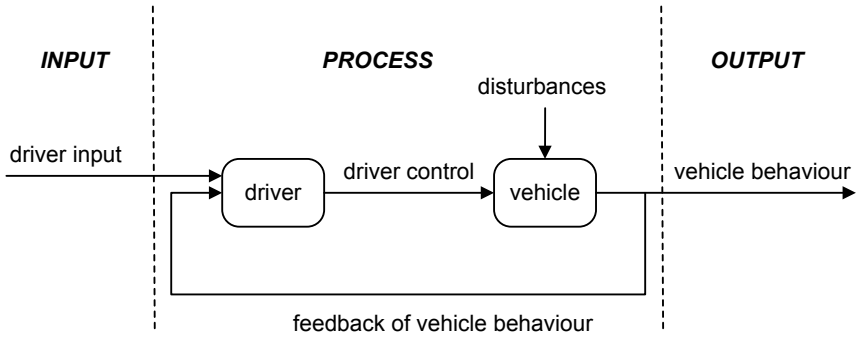


Figure 14. Closed-loop assessment.

2.3.2 Subjective versus Objective Assessment

Handling assessment can also be divided by how the vehicle behavior is assessed, based on the driver's opinion or on measurement data.

Subjective assessment is based on the driver's opinion about various vehicle handling aspects during handling maneuvers. His evaluation of the vehicle's handling behavior can be captured using questionnaires, statements or absolute or relative rating scales. The driver’s evaluation is not only determined by the

vehicle handling behavior, but also by his perception of this handling (Fig. 15). Because of this dependency on the driver, or subject, this is referred to as 'subjective assessment'. In this figure, the driver's subjective evaluation is converted into numerical data by letting the driver score various handling aspects. As these handling tests require excellent driving skills next to the ability to distinguish small differences in vehicle behavior, almost exclusively professional test drivers are used for these subjective assessments. In addition, professional test drivers are able to perform consistently, resulting in reproducible and comparable assessments. Subjective assessments are often used as the final sign-off test for vehicle and vehicle components. These evaluations remain one of the most reliable methods of obtaining good handling quality during design and development (Chen & Crolla, 1998). Nevertheless, no standard for subjective handling assessment exists, neither is there common agreement on how to perform these assessments or how to interpret the results. Although own standardized methods exists within manufacturing companies, sharing of these methods outside the company is not customary, due to confidentiality.

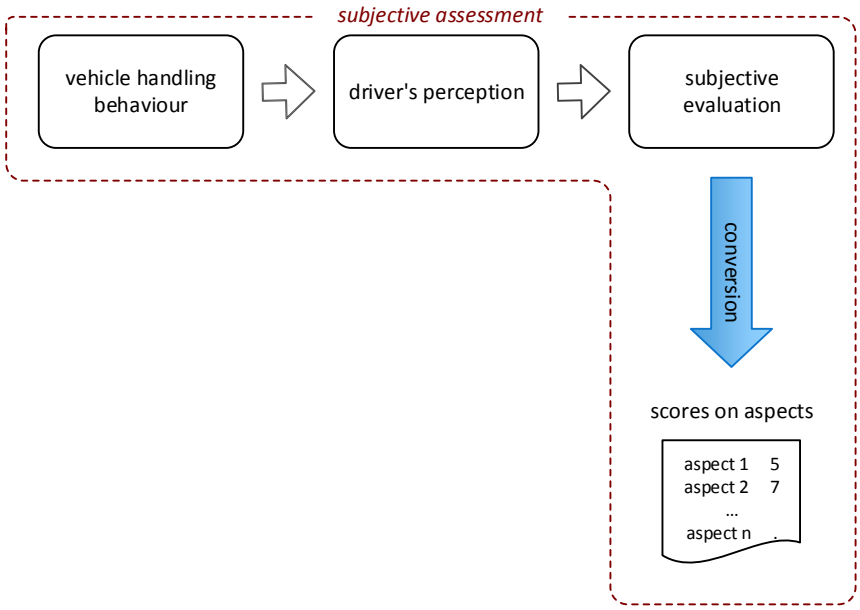


Figure 15. Subjective assessment.

Objective assessment is based on measurements from instrumented vehicles taken during handling maneuvers. The main approach taken is to collect vehicle dynamics measurements during testing, as (transversal and angular) velocities and accelerations in the various directions, often supplemented with other measurements like steering wheel angle and moment data. From these measurements, characteristic data are derived, like delay times (*e.g.*, between

steering input and lateral acceleration), peak values (*e.g.*, maximum acceleration), average values (*e.g.*, speed) (Chen & Crolla, 1998; Dixon, 1996; Heissing & Ersoy, 2011; ISO, 2003; Pauwelussen, 1999b; Sharp, 1999). Fig. 16 shows that objective assessment 'converts' the vehicle handling behavior to these characteristic data. How this conversion is done, what measurements are taken and what characteristic data are derived varies, and choices herein are determined by the goals of the assessment. These characteristic data are then used for calculating the handling criteria values to be compared to preferred (range of) values. Compared to the multi-modal information a driver perceives, the characteristic data derived from the measurements contain less information, which makes it more difficult for objective assessment methods to distinguish small differences in handling behavior, as skilled drivers are able to do. However, even if all the driver information would be available for measurements, it is not yet clear how a driver perceives and processes all this information for his evaluation of the vehicle's handling behavior. This part of the driver is still a black box.

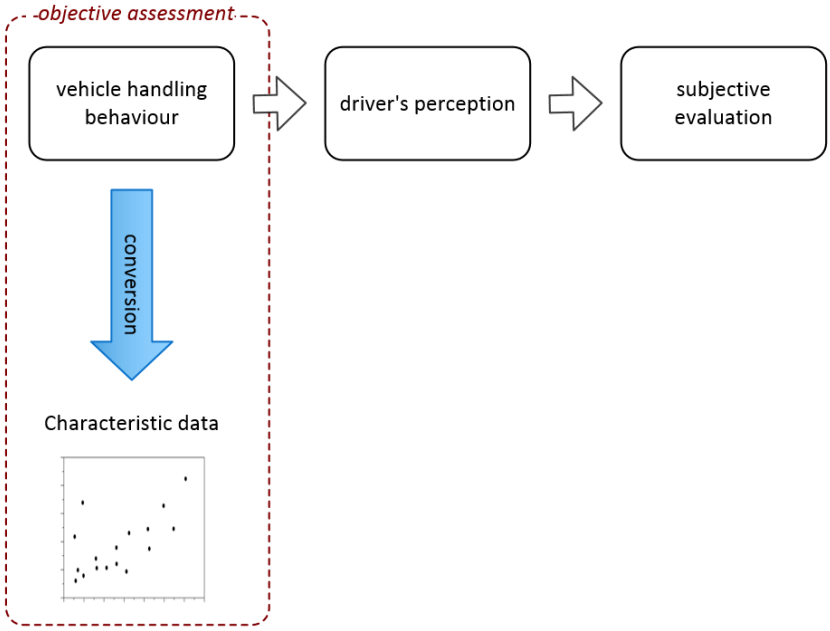


Figure 16. Objective assessment.

The above handling divisions can be combined: open-loop or closed-loop maneuvers can be evaluated objectively and/or subjectively, although the combinations open-loop/objective and closed-loop/subjective are often seen for handling assessments. In several studies (Aurell, Fröjd, & Nordmark, 2000; Bergman, 1973; Chen, Crolla, Alstead, & Whitehead, 1996; Crolla, Chen, Whitehead, & Alstead, 1998; Hill, 1987; Jaksch, 1977; Lincke, Richter, &

Schmidt, 1973; Seghers, 2006; Sharp, 1999; Van Randwijk, Godthelp, Käppler, & Ruys, 1991; Vos, de, Godthelp, & Käppler, 1999), objective measures derived from measurements are correlated to subjective evaluations to get more understanding of important factors influencing handling. This is discussed in more detail in section 3.1.

Several handling assessments are standardized (ISO, 1999, 2003, 2004). These standards all emphasize that the driver-vehicle system forms a closed-loop system with significant interaction, which makes evaluating the dynamic behavior therefore difficult. They state explicitly that the results of these tests can only be considered significant for a small part of the overall dynamic behavior.

2.3.3 Real versus Virtual Testing

To reduce development time and costs, assessment of the impact of component characteristics on vehicle handling performance at an early stage of development can be done using virtual testing. In virtual testing, mathematical models are used for simulating the vehicle, its components and its relevant inputs according to open-loop test procedures. In general, the models used are physical models, based on first principles, as with multi body modeling, with model complexity depending on the model objective. The model parameters are either measured or estimated and additional tuning is often done during validation of the model with real test results. Using such a vehicle model, the influence of tires characteristics can be assessed by simulating the same maneuver with different tire model parameters, which represent different tire characteristics, and comparing the results.

For simulating closed-loop test procedures, a mathematical model of the driver is required. For handling tests, these driver models are often based on control theory; this topic is further explained in section 3.2. By using virtual testing, the combined component-vehicle behavior may be optimized largely, before the vehicle is actually built. In addition, it can derive understanding of the handling behavior by analyzing the model and the model parameters. For example, changing a parameter of the model, like the cornering stiffness of the front tires and subsequently analyzing the resulting change in handling performance can indicate the relevance of this parameter for handling. For simple, linear models, analytical analysis can be performed and linear analyses tools like frequency response assessment can be utilized for the handling behavior. For more complex, nonlinear models, like for modeling high-speed maneuvers, qualitative methods exist and numeric methods are necessary for simulation if quantitative output is required.

Additional advantages of virtual testing are the fact that the test conditions can be kept constant between tests and disturbances can be eliminated (like changing weather). The speed at which various tests can be performed can be easily adjusted, just as other parameters, only limited by the complexity of the models and the processing speed of the computer. In addition, virtual testing can be used for tests that could perform a risk for the driver or else could preferable not be performed in real life.

As models are a simplification of reality, they can never replace all real life testing. Actual testing will always be used for assessment of handling behavior, including all real world nonlinear behavior and influences. Although correlation between objective, simulated and subjective assessment performance is and has been the subject of many studies (see section 3.1), the final subjective handling assessment can only be done with real testing. Virtual testing can therefore be seen as an additional tool for vehicle handling assessment.

2.3.4 Tire Handling Assessment

Objective, open-loop performance assessment of tires is done using tire testing equipment in the laboratory, outdoors with special test equipment or by using virtual tire testing. Tire characteristics related to handling, like lateral force as a function of tire load, slip angle, camber angle can be derived directly. Current practice for the assessment of tire handling is subjective: a vehicle driven on a track by a skilled and experienced test driver. Like with handling assessment, the test driver performs open-loop and closed-loop maneuvers and based on these tests, rates the handling behavior. This is repeated with different tires while keeping all other conditions (vehicle, driver, environment (road), and maneuver) as constant as possible, relating the handling differences directly to tire differences. The restricted time frame for performing the test are for tire handling tests even stricter than for handling tests, due to the major impact of the tire road interface, which is strongly influenced by the weather. The subjective handling assessments are a decisive factor in the final tire handling qualification.

To use subjective assessment results, not only as a final qualification, but also for improvement of tire handling, these results must be converted to requirements for tire performance characteristics. This step is shown in Fig. 5 on page 12 with the arrow going from tire and vehicle handling assessment to tire characteristics. The feasibility of this conversion depends entirely on the skills of the test driver in explaining and quantifying his feeling and, additionally, converting this to tire performance. This knowledge of the 'feel' of the driver is necessary for the tire manufacturer to improve tire handling with subjective assessments. This knowledge is necessary, but not sufficient. Again referring to Fig. 5, the arrow representing the requirements for the tire components properties to obtain the desirable tire characteristics is, as stated in section 2.2.2, also a complex task.

Pauwelussen gave an outline of some previous studies (Pauwelussen, 1999a) where tire parameters were divided in construction parameters like materials used and tire dimensions, service parameters like tire pressure and temperature, performance parameters (which are referred to as 'tire characteristics' in this research) and ageing parameters. The results of these studies indicate that the investigated tire parameters influence vehicle performance assessment through the tire characteristics (performance parameters), as shown in Fig. 17.

Figure 17. Influence of tire parameters through tire characteristics on tire and vehicle handling assessment.

The scope of the present research, to explore the influences of tire characteristics on handling assessment, has also been indicated in Fig. 17.

Changing tire model parameters, like the cornering stiffness, in virtual testing can give information on the influence of these parameters on vehicle handling assessment(Pauwelussen, 1999a). Virtual testing can also be used to tune tire characteristics for optimal vehicle handling behavior (Hendrikx, 1997).

2.4 The Adapting Driver

As handling is determined by the total closed-loop driver-vehicle system, the interaction between vehicle and driver constitutes a significant part of it. For the assessment of global handling objectives like (open-loop) vehicle stability, fast response or understeer characteristics, open-loop vehicle handling tests are important. However, these tests are often not sufficient to distinguish between differences in the detailed design of the vehicle, including the variation in tire characteristics. For this, the driver must be taken into account. This is done by using closed-loop handling tests in combination with subjective assessments where the driver gives his opinion based on his perceived driver-vehicle interaction. In this section, a closer look is taken at this driver-vehicle interaction, especially the fact that a driver can adapt to different vehicle behavior. First, modeling human behavior is described, after this workload is explained, being an important aspect reflecting adapting behavior of a human.

2.4.1 The Quasi-Linear Model

Research into modeling human behavior as a closed-loop controller of a dynamic system started in the late fifties where Tustin (Tustin, 1947) published a first article describing a mathematical human controller model. This was followed by research dominated by the field of (military) flight control engineering, where pilot behavior was modelled based on linear control theory. This early research was summarized in an article by McRuer et.al (McRuer, Graham, Krendel, & Reisener Jr, 1965). Although human control includes nonlinear behavior, like having finite thresholds for input signals or having learning abilities, within certain contexts, human control is most characterized with linear control theory (Jagacinski & Flach, 2003; Jürgensohn, 2007). There were two main reasons for modeling human control behavior with linear control theory (Jagacinski & Flach, 2003). First, this made it possible to predict the total closed-loop system behavior of the human controller and the controlled

dynamic system. Especially the possibility to predict the closed-loop stability of a pilot aircraft combination motivated much of the aforementioned early research in the flight control engineering. The theory of modeling dynamic systems with control theory was readily available and modeling the human control behavior using the same control language made it possible to perform this closed-loop analyses. The second reason for this human control modeling was that this could provide additional information about basic properties of human control behavior by analyzing the resulting quantitative models of the human controller.

One of the first human control models is the quasi-linear model (McRuer & Krendel, 1959a, 1959b), described in the Laplace domain as:

$$H(s) = \left(\frac{e^{-\tau s}}{\tau_N s + 1}\right) \cdot \left(\frac{K(\tau_{lead} s + 1)}{\tau_{lag} s + 1}\right) \tag{9}$$

with describing function $H(s)$ for the human controller. The term in the first parenthesis, represent inherent limitations for the human controller, a reaction time delay τ and neuromuscular lag time constant τ_N. Although these parameters can vary depending on different driver characteristics, like age group, level of skill or attentiveness, often these parameters are taken as constants when a certain driver category is modelled. Typical values for τ lie between 0.1 and 0.2 s (Jagacinski & Flach, 2003; McRuer et al., 1965) and for τ_N between 0.03 and 0.10 s (Cavanagh & Komi, 1979). The term in the second parenthesis, represent the adaptive behavior of the human controller, consisting of a gain K to represent proportional behavior, a lag time constant τ_{lag} to represent first order or integrator behavior and a lead time constant τ_{lead} to represent lead or differentiator behavior. This last behavior is often referred to as anticipating behavior, the human controller anticipates on future input signal behavior by taking the current change of input signal into account. The actual parameter values depend on the dynamic system to control and the type of input signal. Together with the term $e^{-\tau s}$, these aspects deliver the model its prefix 'quasi', it is only linear in the considered context (Jürgensohn, 2007). This driver model is until today often used (Plöchl & Edelmann, 2007).

2.4.2 The McRuer Crossover Model

Although the quasi-linear model could represent human control behavior fairly well, for different dynamic systems to control, researchers found different parameter values for the human controller, so a general human controller model could not be found. McRuer and his fellow researchers (McRuer et al., 1965; McRuer & Krendel, 1959a, 1959b) found that a general model could be identified when considering human and system together. This well-known 'crossover model' is defined as the open-loop transfer function of human controller and dynamic system in the region of the gain crossover frequency, and reads in the Laplace domain:

$$H(s)G(s) = \frac{\omega_c e^{-\tau_e s}}{s} \tag{10}$$

with describing function $H(s)$ for the human controller, transfer function $G(s)$ for the controlled dynamic system, ω_c for the crossover frequency and τ_e for the delay time (which is not identical to the human delay time τ from Eq. (9), but is a catchall term that serves to incorporate phase effects near the cross over frequency). The adaptive nature of a human controller is recognized here. This model predicts that the human controller will adapt his behavior to the dynamics of the controlled system to keep the overall (closed-loop) system behavior stable and approximately constant. For a closed-loop system to be stable the open-loop transfer function should have positive gain and phase margins and this is what a human controller tries to accomplish by adapting his behavior. From basic control theory, it can be derived that this open-loop gain ω_c is equal to the bandwidth of the closed-loop system. The human controller will try to adapt this value to get good closed-loop system performance: high enough so that the output can follow the input well, but low enough not to become near instability. This results in a certain crossover frequency and time delay of the crossover model, which is depending on both system dynamics and human characteristics.

This crossover model explained why researchers found different transfer functions for human control behavior for different controlled systems. The crossover model is only stated valid in the region of the gain crossover frequency, which is often the region of interest when considering closed-loop performance. But as variations in the open-loop describing function imply much less variations in the closed-loop describing function in the region below or above the crossover frequency, the model is used often applicable in these regions (Jagacinski & Flach, 2003). For attentive automobile drivers in steering regulation tasks the crossover frequency is about 4 rad/s (Macadam, 2003). This is in line with values found for this research for the double lane change driven by professional test drivers with different tires (described in section 4.6), which are around 3.5 rad/s (Oort, van, 2011). Although the crossover model has only two parameters, it is a very precise description used until today for many human control situations (Jürgensohn, 2007; McRuer et al., 1965; Pronk et al., 2000).

While humans can adapt to different system dynamics, they have limitations that are reflected in the variation of transfer functions that can be used for human control behavior using the quasi-linear model. As this model is a second order transfer function and the crossover model states that the open-loop transfer function resembles an integrator with dead time, this theory defines the maximum order of the system which a human can control is second order (Pronk et al., 2000). If the lag time constant is small, so human integrator behavior only at high frequencies, and the lead term time constant is significant, so lead or anticipating behavior for lower frequencies, the crossover model can be formed for the open-loop transfer function near the crossover frequency. This is in accordance with actual human control behavior found by several researchers and summarized in (Jagacinski & Flach, 2003). Humans can

control zero or first order dynamic systems fairly easy, while this requires only use of the gain and lag time constant, requiring little effort. An example is controlling a cursor on a computer screen with a computer mouse (zero order system) or controlling the opening or closing of a car window with a button (first order system). Without almost any practice, humans can control these. For control of second order systems the lead term time constant must be used in combination with no lag time constant, the human must act as a differentiator and this requires considerable more effort, although most humans can learn this. An example is steering a car, a car has dominant second order behavior if we consider the steering wheel angle as input and the position on the road as output of the system. Most people are able to learn this, with some practice. Higher order systems are very difficult for humans to control without help or extensive training and not all humans are capable of learning this, like backing up a truck with trailer.

This crossover model also explains why overall, closed-loop system performance is not always a good measure for the performance of a dynamic system controlled by a human, due to his adapting behavior. This adapting effect is also found in handling, for example in a study (Vos, de et al., 1999) where vehicle handling conditions were manipulated by varying rear tire pressure, resulting in different vehicle understeer behavior and lane width. Despite these varying handling conditions, drivers were able to maintain a constant performance level by adapting their control behavior, in this case steering effort. Subjective assessment revealed higher ratings for lower steering effort, which indicates a relationship between the amount of adaptation required to keep the performance up, the effort or workload and the subjective rating. This example shows a common known close relationship between adapting behavior of a driver, his willingness to invest effort and his perceived workload. These aspects and their relationship with overall performance are therefore explained in the next sections.

2.4.3 Workload

Workload can be divided into physical and mental workload. Physical workload is related to the work muscles do (O'Donnell & Eggemeier, 1986) and in the context of driving it can be measured by movements done by the driver. From steering wheel measurements, steering measures can be derived, like steering reversal rate in the time domain or measures based on the frequency components in the frequency domain, like the High Frequency Area (HFA). Because these measures are based on movements of the driver, they can be regarded as physical workload measures, but due to the strong correlation with mental workload, they are often considered as mental workload measures (MacDonald & Hoffmann, 1980; Pauwelussen & Pauwelussen, 2004; Verwey & Veltman, 1996). Mental workload is more complicated to define and measure. It is a general term, which lacks a clear overall definition. Instead, it has various meanings and definitions in different contexts (Hacker, 2006; Megaw, 2005). A definition used for studying driver's mental workload is given by De Waard (Waard, de, 1996) as 'the specification of the amount of information processing

capacity that is used for task performance'. Information processing capacity is the mental capacity including not only cognition but also motivation, including a driver's willingness to put in effort for reaching task performance. Mental workload does not depend on the task demand directly, but is influenced by how the task is perceived by the driver (Megaw, 2005). This indirect relation, together with other related concepts, is sketched in Fig. 18 and explained below.

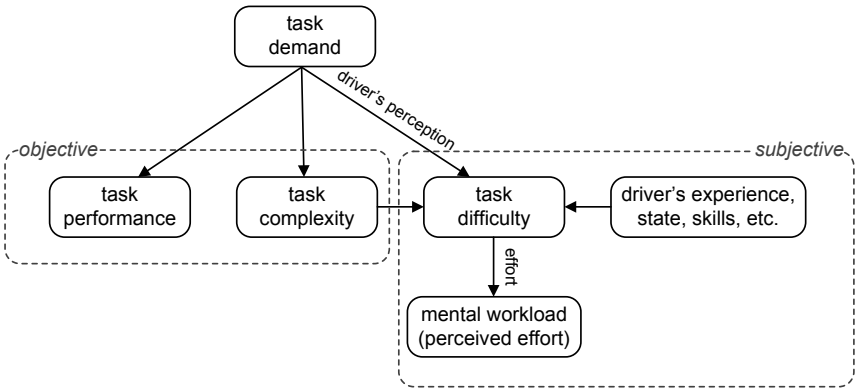


Figure 18. Relation between task demand and mental workload.

Task performance is an objective quantification of how well a task is done, measured mostly in terms of accuracy and speed. An objective description of the task demand is task complexity, being a function of the required processing stages to perform that task and external situational factors like visibility. The driver's perception of the task demand defines the task difficulty for him, together with the task complexity and his experience, state (e.g., fatigue, stress), skill, motivation, *etc.* This makes task difficulty a subjective property of task demand. For example, the task demand for an experienced and a novice driver is to drive a lane change marked with cones with a predefined speed of 80 kilometers per hour, without hitting one of the cones. The task performance can be measured objectively for both drivers by the number of cones that is hit. The task complexity is also the same for them; these are the actions that must be performed, like steering, looking ahead and anticipating the vehicle behavior. The difficulty of the task is subjective: this will be less for the experienced driver than for the novice driver. Driver's mental workload is determined by this task difficulty by the amount of effort a driver is willing and able to invest for task performance. Task difficulty and subsequently mental workload is therefore a subjective property. In this example it is possible that both the experienced and the novice driver have the same performance, *e.g.*, no cones hit during the lane change, although the mental workload of the novice driver is much higher than for the experienced driver because the novice driver must invest more effort to reach this performance.

2.4.4 Performance and Workload

A model that relates task performance and mental workload to task demand is given by De Waard (Waard, de, 1996). The part of the model that is related to medium and high task demand is reproduced in Fig. 19. The original model also constitutes a part for the low demand region. This low demand part is not relevant in this research, because test drivers are not subjected to low demand during handling testing and is therefore left out. For the medium to high demand part of the model, four regions are defined. Region $A2$ indicates high performance and low workload, even with increasing demand. With further increasing demand in region $A2$, the performance is unaffected, but therefore, the driver must invest a certain amount of effort depending on the task demand - the so-called task related effort - to keep this performance up. In region B, the increase of demand is accompanied by a decrease of performance together with further increase of workload. In this region, the driver cannot keep up the performance, despite his effort. In region C, the demand is excessive and leads to very high workload and very poor performance, the driver is overloaded and cannot meet the task demand at all. Although task demand is objective, the positioning of the regions on this 'task demand axis' is subjective. It depends on the perceived difficulty of the task, which depends on the driver as is explained and shown in Fig. 18.

Of special interest for this research is the region $A3$ that is shaded in the figure. This region corresponds to the adapting driver theory. While performance measures will show no changes with increasing demand in region $A3$, the driver will adapt to this increasing demand by investing more effort, increasing his workload, to keep the performance at the same level. This increase in workload as perceived by the driver can influence his subjective evaluation. Further increasing of demand causes the transition to region B, which is indicated by performance decrease.

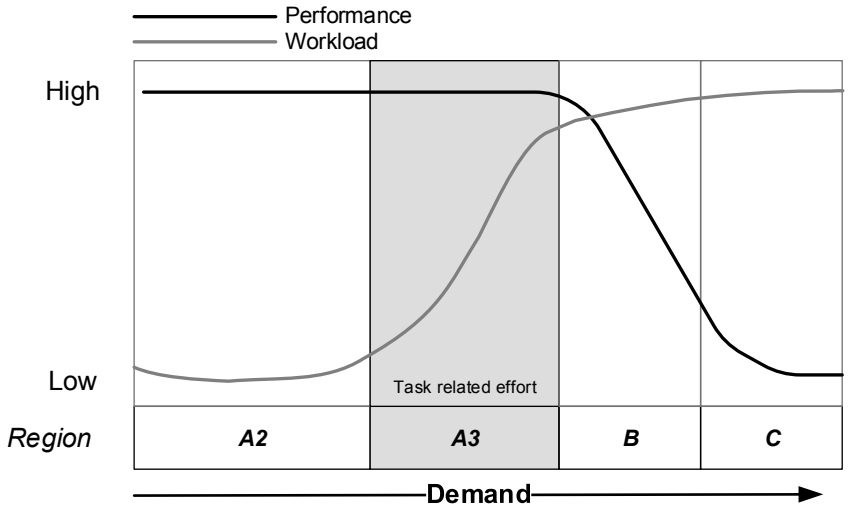


Figure 19. Task performance and workload as a function of demand (Waard, de, 1996).

During testing of tires with small differences, which can be assessed by a professional driver but cannot be shown by performance measures, assumptions made in this research are, that the test driver is acting in region *A*3 and that the amount of workload or task related effort is correlated with the subjective evaluation (for validation of assumptions, see section 5.1 and 5.3). These assumptions are made because tire test drivers, having the explicit instruction to differentiate between different tire behaviors, will often push the tires to its limits by performing high speed maneuvering to be able to identify those differences. This corresponds to exploring the positions of the regions *A*3 and *B* for a specific tire. For example, assume a test driver to evaluate the handling behavior of different tires by performing a lane change maneuver at a fixed high speed without hitting one of the cones. This instruction forces the driver to the high demand, high performance region *A*3. With the best handling tire the driver is able to keep the performance high, with relatively low effort, so low workload (left side of region *A*3) while with the less good handling tire the driver must invest much effort to keep the performance high (right side of region *A*3) or possibly is not able to keep the performance that high, so hitting a cone (region *B*). The assumption is that higher workload causes a lower subjective evaluation, because the driver is forced to invest more task related effort (higher workload) to keep the performance at a high level (for validation of assumptions, see section 5.1 and 5.3). Therefore, this is the region of interest for this research.

2.4.5 Measuring Mental Workload

For measuring mental workload, measurements techniques can be divided into three main categories (O'Donnell & Eggemeier, 1986): subjective measures, performance-based measures and physiological measures.

Subjective measures are evaluations of the perceived workload done by the operator himself. Often used for these measures are rating scales, for example with ratings from 1 to 10 or from 'very bad' to 'very good'.

Performance based measures can be split in primary and secondary task performance measures. Primary task performance measures are based on how well the operator performs on the main task demand put upon him, mostly based on accuracy and speed. For the task "driving a certain maneuver at the highest possible speed without hitting any cones", the primary performance measure could be the highest possible speed. During primary task performance, an additional, secondary task can be placed upon the driver, like reacting to a light or audio signal and measuring the reaction time as secondary performance measure or giving additional tasks, like counting backwards or do calculations and use errors as performance measure. For a secondary task, it is important to instruct the operator well to give the primary task priority. Performance measures of both primary and secondary tasks can give an indication of the workload necessary for the primary task.

Physiological measures are based on the assumption that changes in workload are reflected by changes in the level of arousal of the operator, which changes the activity of his nervous system. Examples are measurements of heart rate, respiratory, brain activity and muscle activity.

For measuring mental workload, the most important property of a measure is the ability to distinguish different levels of workload, which is called 'sensitivity' of a measure. Sensitivity must always be accompanied by the regions of performance where this sensitivity is valid: outside these regions, the measure cannot give an indication of workload. Subjective measurements are frequently used for assessing workload in the task related effort region *A*3. As workload is a subjective property, assumption is made that the subject himself can best assess this, although it can be difficult for drivers to distinguish between physical and mental workload. Clearly, primary performance measures are not sensitive in region *A*3, because performance does not change in this region, but they are sensitive in region *B* and *C*, because in these regions performance does change. So performance measures can be used to indicate the transformation from region *A*3 to *B* with increasing demand. From the physiological measurements, only a few can be used outside the laboratory and in a vehicle environment. Heart rate measure, more precisely, frequency analyses of the heart rate variability (HRV) is one of the best candidates (Waard, de, 1996) and will be explained hereafter. For time domain analyses, heart rate data can be used. Increase of mental workload will increase the heart rate resulting in lower values of the interval between heartbeats, the Inter Beat Interval (IBI). Using heart rate or IBI in the time domain as only indication of mental workload is not distinctive enough, because heart rate increase can have more causes, physical workload being a dominant one (Kamalakannan, Groves, & Freivalds, 2007). HRV is the variability of this IBI-data. In the frequency domain, spectral analyses of this HRV has shown that the power in the band around the 0.10 Hz component of the HRV is related to mental workload if the duration of the workload is not too short, *i.e.,* longer than 30 seconds (Mulder, 1992; O'Donnell

& Eggemeier, 1986; Stassen, Johanssen, & Moray, 1990; Verwey & Veltman, 1996; Waard, de, 1996), especially for task related effort, with the advantage that other influences, like physical workload, are small. A distinction can be made between the power in a low frequency (LF) and high frequency (HF) band. LF range is defined as 0.04-0.15 Hz and HF range as 0.15-0.4 Hz. An increase of the LF component and decrease of the HF-component can indicate mental workload (Guger et al., 2004). The ratio of LF/HF can be used as an indicator for mental workload for drivers (Monsma & Sultania, 2011).

Although task performance measures during secondary task demands are also sensitive in region *A*3 and can be used well outside the laboratory, difficulties are often encountered between the imposed priority of the primary and secondary task and the actual priority of the driver. A solution is to use embedded secondary tasks, which is a task performed during normal operation of the primary task, but which can be assessed separately (Waard, de, 1996).

3. Research Approach

'Understanding the driver' in this research means determining *what* handling behavior the driver considers as 'good', as explained in section 2.1.2, but also, *how* this handling behavior influences his perception and subsequently his subjective evaluation. This is especially important owing to the adapting behavior of the driver, which can result in tests with similar objective handling behavior but with different subjective assessments. Is the driver forced to adapt much to the vehicle behavior? Does he experience a high workload? Is he required to put in much effort or is it relatively easy to perform well? Does he have to anticipate much? Getting answers to these perception related questions will provide this understanding of the driver, which can contribute to developing well handling and therefore safe vehicles.

In this chapter, the three methods for understanding the driver that are followed in this research, including background information, are described. The first two methods relate objective to subjective measures to provide information about *what* handling behavior is considered as good handling behavior. The third method models the driver in the closed-loop relationship with the vehicle to provide information about the perception of the driver to clarify *how* this handling behavior influences the driver.

3.1 Answering the 'What'- Question: Relating Objective to Subjective Measures

Seen from a system modeling view, the vehicle handling behavior could be considered as the input to the driver. The driver perceives this input and makes his assessment, which is the process of interest. This process results in the driver's subjective evaluation of this handling behavior, which can be defined as the output. This is shown in the top part of Fig. 20. As explained in section 2.3, objective assessment methods convert handling behavior through vehicle dynamics measurements to characteristic data (left-hand side of Fig. 20). Subjective assessments methods convert the driver's opinion to data, like scores on handling aspects (right-hand side of Fig. 20).

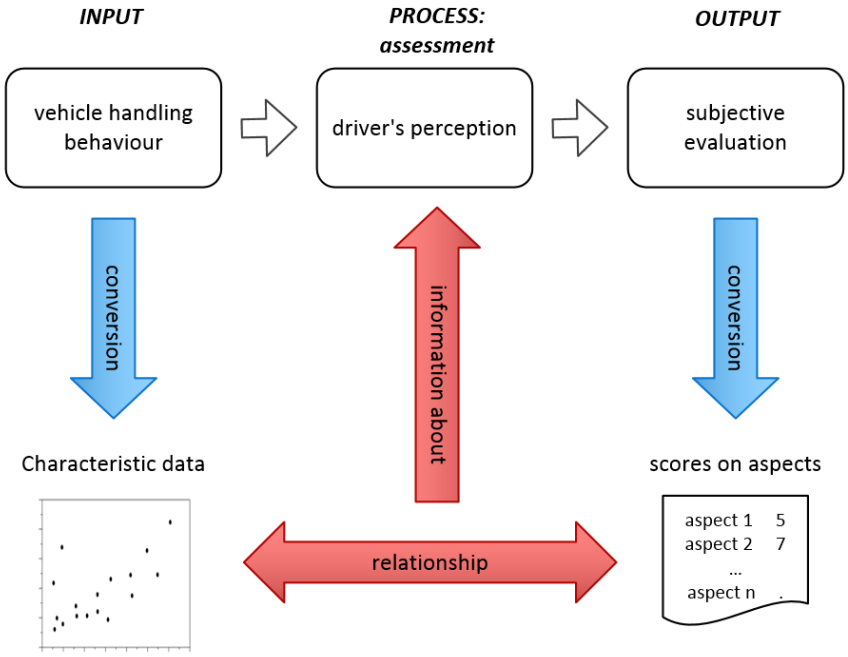


Figure 20. Relating objective to subjective measures to get 'what'-information about the driver's perception.

Studying the relationship between the data from objective and subjective assessment is a frequently used method to obtain information about the driver's assessment process (middle part of Fig. 20). Such studies (Aurell et al., 2000; Bergman, 1973; Chen et al., 1996; Crolla et al., 1998; Hill, 1987; Jaksch, 1977; Lincke et al., 1973; Seghers, 2006; Sharp, 1999; Van Randwijk et al., 1991; Vos, de et al., 1999) often use correlation or regression methods for analyzing this relation. Usually, the set of characteristic data coming from the vehicle dynamics measurements is large and not independent, making statistical analyses difficult. Therefore, this set is often reduced by using factor analyses or principal component analyses and sometimes combined with additional clustering techniques. Although using these reduction techniques can facilitate finding relations, their application is not standardized and requires significant knowledge. Additionally, there is no clear interpretation of the results. Still, these studies of objective-subjective relationship of handling have resulted in global indicators for good handling, as presented in section 2.1.2.

Next to using different methods for relating the objective data to subjective data, also the conversions from real life to numerical data differ for the approaches taken in the past. This conversion largely determines the amount of information that can be gathered by studying the relation. For example, if no steering wheel data are measured during objective assessment, the information about the important steering wheel feel of a driver is strongly restricted. This also applies for the subjective data, if the steering wheel feel is not part of a

handling aspect that is rated by the driver, this information will not show up by analyzing the relation. All these differences in these objective-subjective approaches make comparison and overall conclusions difficult. Sharp (Sharp, 1999) evaluates several of these approaches (Bergman, 1973; Crolla et al., 1998; Jaksch, 1977; Lincke et al., 1973), concluding that although correlations are found, the results are in most cases not applicable outside the context in which they were derived. Additional, there is no clear interpretation and standardization of this relation. He states that the only exceptions are objective indicators that have a sound fundamental basis of understanding for them (see section 2.1.2). Regarding to tire handling assessment, Pauwelussen (Pauwelussen, 1999a) evaluated several studies which included objective-subjective relations. He came to similar conclusions. He also found a lack of clear interpretation and standardization for this relation. Also in these studies, there is a general understanding that there are objective indicators that are important for subjective evaluation (see section 2.1.2).

To bypass some of the limitations mentioned here for this research some other approaches were followed, as mentioned at the end of section 1.1.

In the first method described in chapter 6, the driver's subjective handling assessment is predicted based on vehicle dynamics measurements with a General Regression Neural Network (GRNN). By using a GRNN for this, the previously mentioned limitations for using regression can be circumvented (see section 6.2 for explanation and references).

The second method described in chapter 7, focuses on the driver, more specifically, on his workload as an indication for the subjective assessment. This derives from the fact that the driver adapts to changing vehicle handling behavior (see section 2.4 for explanation and references). The assumption made and validated in this research is, that the more driver adaptation is required, the more workload the driver perceives, the lower his vehicle handling assessment score will be. Because workload cannot be measured directly, indirect measures must be used. This includes both physical as mental workload measures.

3.2 Answering the ‘How’-Question: Driver Model Method

Although studying the relationship between objective-subjective data is able to provide information on what the influence of vehicle's handling behavior on driver's evaluation is, it does not give much information on how this influences the driver's evaluation. To obtain this information a driver model method is derived in chapter 8. The basic idea is not to model the driver's assessment process as a high level open-loop input-process-output system, as done in the other methods and shown in Fig. 20, but to go one step deeper and to model and simulate the driver in its closed-loop connection with the vehicle, as shown in Fig. 21. Based on the adaptive nature of a human controller (McRuer & Krendel, 1962; McRuer & Weir, 1969), human control parameters can reflect control behavior, but also different levels of workload (Jagacinski & Flach, 2003), and can be related to subjective evaluation (Abe, Kano, & Shibahata, 2009). This inspired this research for the third method. It focuses on the driver but not by

looking at measures from 'outside' the driver, like workload measures, but by modeling the driver behavior and looking at driver parameters 'inside' the driver (model), as shown in Fig. 21.

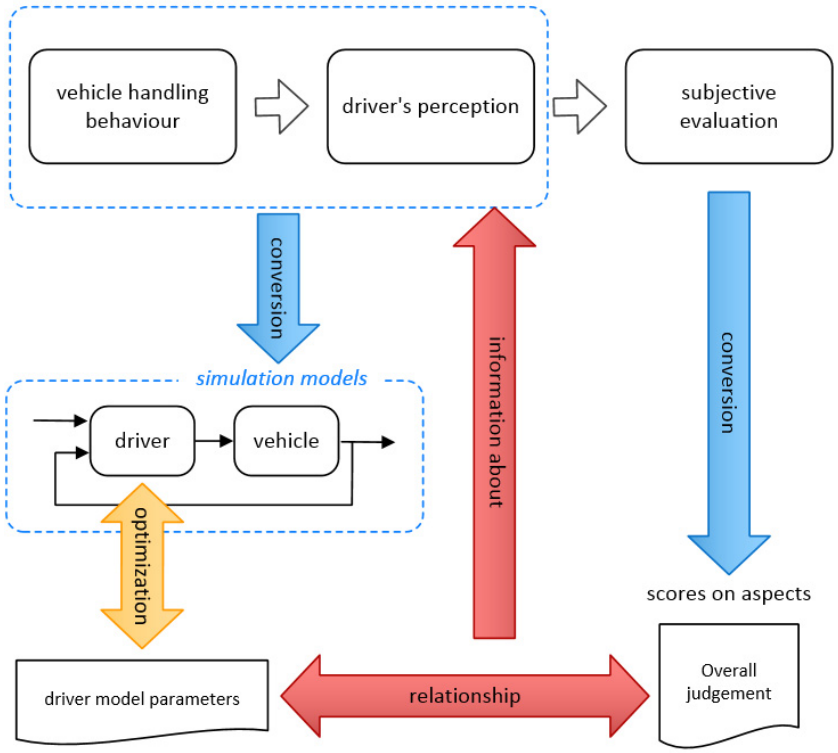


Figure 21. Driver model method: relating driver model parameters to subjective evaluation scores to get 'how'-information about the driver's perception.

For this method, models of the vehicle including tires are run in closed-loop with a driver model, to simulate a handling maneuver. If the driver is modelled with parameters that are related to certain driver's characteristics, like effort or workload, these could relate to the subjective assessment of tires. If different tires (tire models) result in different (fitted) driver parameters, these 'tire dependent driver parameters' could provide information about the amount of effort and/or workload a driver perceives. This, in turn, could give an indication of the subjective evaluation of the tire handling behavior.

4. Field Experiments

Several field experiments were performed for this research and are described in this chapter. The following chapters, describing the three different methods regarding driver's handling assessment, will refer to this chapter for the relevant experiments. Before describing the field experiments, the general test method, the drivers, the selected tires, the test vehicle, equipment and the test tracks, are described.

4.1 General Test Method

For all experiments performed for this research the ISO standard for vehicle dynamics test methods regarding passenger cars (ISO, 2006) was followed for variables, measuring equipment, data processing, test track, test vehicle preparation, initial driving, and general data and test conditions. For the mental workload test, the test track was not dry but had changing humidity, which could not be avoided. This effect is taken into account in the testing itself by adapting the task demand and for the modeling part, the vehicle model is adapted to the changing circumstance (further explained in the relevant sections). All tires were set at the tire pressure advised by the tire manufacturer and were run in mounted on the test vehicle for at least 150 km with normal driving on public roads. The position of the tire on the vehicle was registered and kept the same during run in and testing. Between run in and actual testing the tires were kept in regulated conditions. Before testing, tires were warmed up to normal driving conditions. During all testing the vehicle drove in third gear with the electronic stability control completely turned off, so it could not interfere with the driver actions. Before the actual testing of a tire by a driver, this driver drove several maneuvers to become familiar with the new conditions, to prevent learning effects during testing. Between testing of different tires performed by one driver there was always a rest period for that driver (mostly when another driver performed testing) to avoid carry over effects of testing, *i.e.*, where a test could be influences by the previous test. For example, if the test is tiring for the driver, this could influence a next test if no time for the driver to recuperate was given.

4.2 Drivers

As stated in the introduction, the scope of this research is subjective tire handling assessment by professional test drivers. Therefore, two professional tire test drivers of the tire manufacturing company Appollo Vredestein B.V. participated in this research. Due to confidentiality, it was not possible to include professional tire test drivers from other (tire) companies. One test driver was in his thirties and had a few years of experience as professional tire test driver after several years of practice as test driver. The other test driver was in his fifties with over twenty years of experience as professional tire test driver. They performed all subjective tire handling assessments (described in section 4.6) and participated in the mental workload double lane change experiment (described in section 4.7.5).

As mentioned in the introduction, tire subjective testing is also performed by highly skilled test drivers from magazines, as basis for their articles. Therefore, a group of eight highly skilled nonprofessional drivers in the age between 21 and 36 participated in this research. The drivers were selected by the research team from their network based on driving skills often gained through (nonprofessional) rally driving or racing competitions. Prior to the actual tire testing all drivers had extensive practice to familiarize themselves with the vehicle, tires and maneuver. Two drivers participated in the mental workload slalom experiment (described in section 4.7.4), six drivers in the mental workload double lane change experiment (described in section 4.7.5).

4.3 Tires

For this research, six different tires were selected. The tire specifications are given in Table 1.

Table 1. Tires used in experiments.

#	Tire Dimension	Run flat	Type
1	205/55R16	no	Winter tire
2	225/45R17	no	Winter tire
3	225/45R17	yes	Winter tire
4	225/45R17	no	All seasons tire
5	225/40ZR18	no	High performance summer tire
6	225/40ZR18	no	High performance summer tire

These six tires were selected based on their expected handling performance as assumed by the tire manufacturer based on their data, test results and experience. Fig. 22 shows the relative expected handling performance for all six tires.

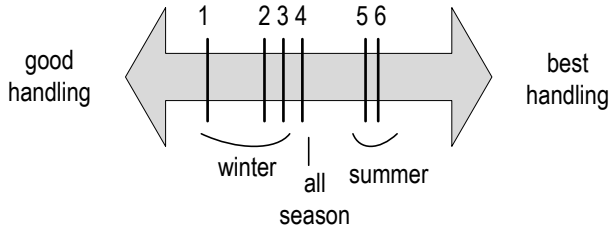


Figure 22. Relative expected handling performance for all six tires.

Tire Models

To obtain the nonlinear Pacejka Magic Formula tire models (see section 2.2.4) for all six tires, the tire characteristics were identified by TNO by performing measurements for pure cornering and pure braking for three different tire loads on two different tire pressures (front and rear) for the tires mounted on the TNO test trailer on a test track. With these measurements, all the relevant Magic Formula coefficients were identified. These coefficients are used with the MF Tyre model of Delft Tyre that implements the Pacejka Magic Formula. This model can also interpolate between parameter values, like changing tire loads during driving, giving realistic values.

Some tire characteristics, plotted from the tire models are shown next. Fig. 23 shows the normalized lateral force μ (Eq. (2)) for three different tire loads for the front tire pressure.

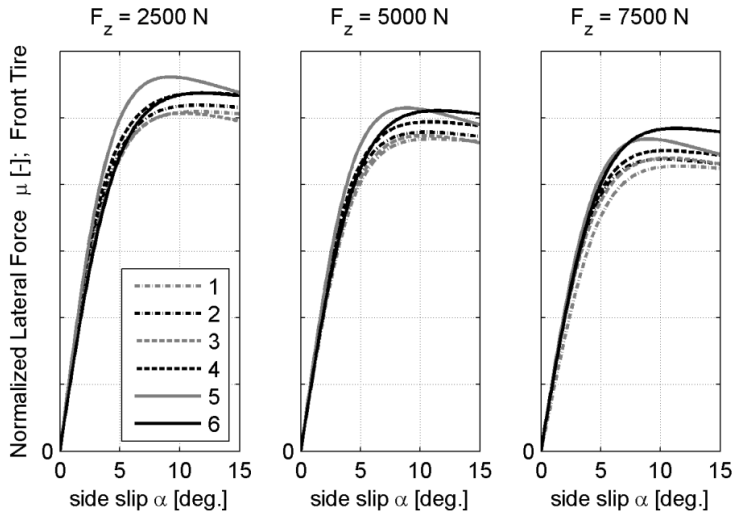


Figure 23. Normalized lateral force μ for all six front tires as function of side slip angle α for three different tire loads F_z.

In agreement with the common tire theory, the normalized lateral force decreases with increasing tire load. As expected for the high performance

summer tires 5 and 6, they produce the highest lateral forces, although for low slip angles and low tire load this is not true for tire 6, combined with lowering cornering stiffness in this case. The lateral force behavior of the two summer tires is rather different. The lateral force of tire 6 is relatively increasing with higher tire loads and higher slip angles compared to the other tires. This implies that tire 6 will perform better during more extreme cornering, having less loss of lateral force due to the weight transfer from inner to outer wheels. This is combined with almost no peak for higher slip angles, as is shown for tire 5, tire 6 keeping the maximum lateral force value almost constant. This gives more predictable, desirable behavior for the driver, having no 'sudden' loss of lateral force when the slip angles increase and the lateral force goes 'over the peak'. Tire 4 shows overall lateral behavior between the winter and summer tires, but for low tire load it has even more lateral force than tire 6. Tire 1 shows the lowest lateral force and most decrease of lateral force due to increase of tire load, which is in agreement with being the less good handling tire.

Fig. 24 and 25 show the pneumatic trail and normalized aligning moment (Eq. (1)) for three different tire loads for the front tire pressure. Fig. 26 shows the contact length for the front tire pressure.

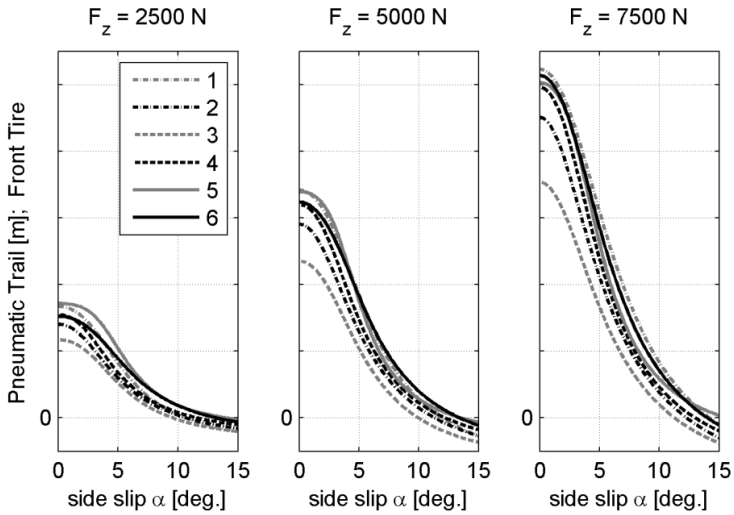


Figure 24. Pneumatic trail t_p for all six front tires as function of side slip angle α for three different tire loads F_z.

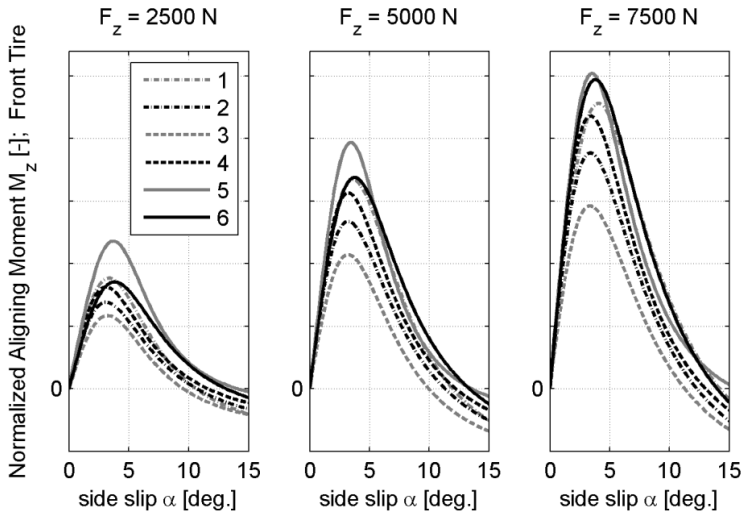


Figure 25. Normalized aligning moment M_z for all six front tires as function of side slip angle α for three different tire loads F_z.

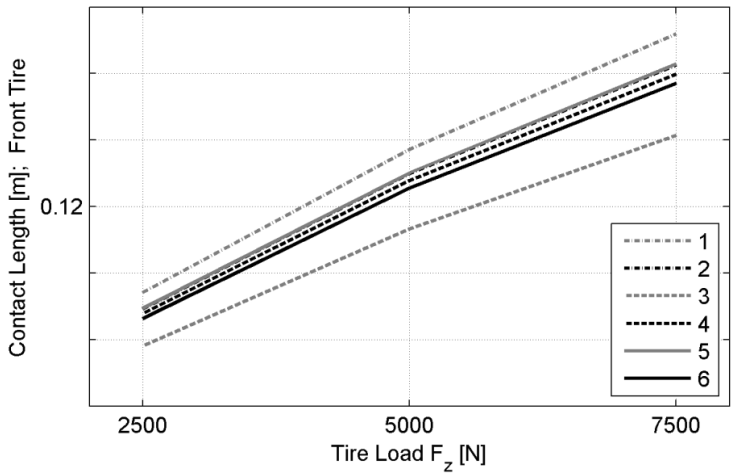


Figure 26. Contact length of tire contact path for all six front tires as function of tire load F_z.

In agreement with the common tire theory, the pneumatic trail increases with increasing tire load. Also for this characteristic, the summer tires have high values, tire 6 catching up with tire 5 for higher tire loads. Tire 1 has a comparable high value, which is due to the fact that this tire has the smallest width, as can be seen in Table 1. This also causes the largest contact length of this tire, as can be seen in Fig. 26. Tire 3 has the lowest values for these characteristics, because this is a run flat tire, having more vertical stiffness of the tire, causing less deformation under load. The value of the aligning moment is determined by the lateral force times the pneumatic trail. This aligning moment is felt by the driver

in the steering wheel and contributes therefore significant to the tire feel, it gives the driver information on the grip of the tire. In addition, for tires with the same width, but lower aspect ratio, the height of the sidewall is smaller, causing less lateral deformation and more direct steering feel, which is in favor of tire 5 and 6.

4.4 Test Vehicle and Equipment

The test vehicle was a BMW 320i Touring equipped with a measurement system logging all vehicle dynamics variables and additional steering wheel angle and moment. The main signals that are measured with the vehicle and used in this research are listed in Table 2. The signals have been sampled with 100 Hz. Video cameras were used in- and outside the vehicle, capturing the driver and the road ahead to be used for analyses and additional synchronization of the data.

Table 2. Main signals measured and used in this research.

Signals	Sensor	Company
Velocities	CORREVIT® HS-CE / H-CE	CORRSYS-DATRON Sensorsysteme Gmbh.
Steering wheel angle, angular velocity and moment	Measurement Steering Wheel	
Yaw velocity	CRS03®	Silicon Sensing Systems Ltd.
Lateral acceleration	SMI model 7130	Silicon Microstructures Inc.
Optic cell		
Position and secondary task control	imc Devices®	imc Meßsysteme GmbH
Heart rate measurements	BioNex 8 slot chassis, Model 50-3711-08 / Biolab software	MindWare technologies
Video	Camcorder DCR-SR35E	Sony

Fig. 27 to 29 show pictures of the test vehicle and the equipment. The mass of the vehicle including driver and equipment was 1615 kg, with front axle static load of 7325 N and rear axle static load of 8825 N.

Figure 27. Front velocity and height sensor and GPS antenna on the test vehicle (left) and main computer and data logger (right).

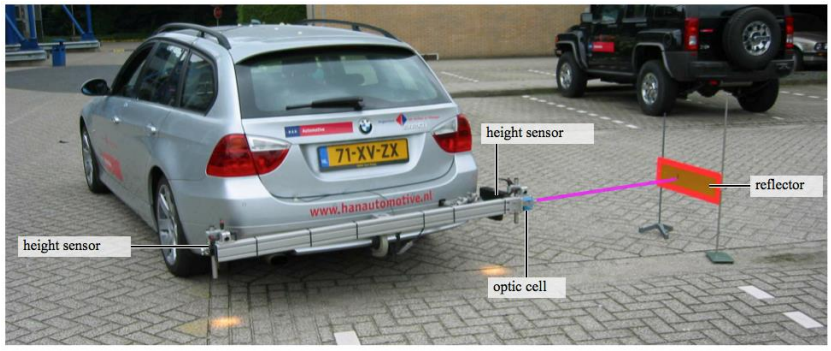


Figure 28. Rear velocity and height sensors and optic cell on the test vehicle and reflector.

Figure 29. Measurement steering wheel, microphone, reaction lights and interfaces of main computer and data logger.

4.5 Test Tracks

Almost all experiments were performed at the RDW-proving ground in Lelystad. The part of the proving ground that is used for this research consists of a steady state circle, a dynamics test area and a straight track to accelerate (Fig. 30).

Figure 30. Part of the RDW-proving ground in Lelystad that was used for the experiments.

The workload experiment double lane change tests of the nonprofessional driver group was held at the former military air force base 'De Peel' in Vredepeel (Fig. 31). The tests were held at the taxi strip of the runway, which has asphalt that is similar to asphalt on public roads.

Figure 31. Part of the air force base in Vredepeel that was used for the experiments.

4.6 Subjective Tire Handling Experiment

The objective of this experiment was to perform subjective tire assessments to collect subjective evaluations of the tires used in this research. For this experiment, two professional tire test drivers participated. Although their regular subjective tire testing was done on the Nürburgring in Germany during extreme driving, for this research an alternative tire testing maneuver was designed, to be held on a proving ground. This was done in close consultation with the drivers to ensure that all relevant tire handling aspects could be assessed. In one part of the driving maneuver, the driver could steer freely with varying speed over part of the proving ground to perform the tire handling assessment. In the other part, a double lane-change maneuver was set up with cones (Fig. 32) through which the drivers drove several times during one test. One test stands for driving the whole maneuver for subjective evaluation of one set of a certain tire by one driver. The offset between the lanes was set to 5.5 m, instead of the 3.5 m prescribed in the ISO double lane change (ISO, 1999), to make the maneuver more challenging for the drivers without having to drive with very high speed. The cones used for marking the maneuver were small, flat cones with a height less than 100 mm instead of the cones stated in the ISO description with a minimum height of 500 mm, because these high cones could damage the test equipment. This made driving the maneuver more difficult for the drivers, because the visibility of the cones was less good. Therefore,

sometimes cones were hit, without the driver noticing. When driving back from the maneuver, drivers were informed by arm movements of the helpers if a cone was hit and which one.

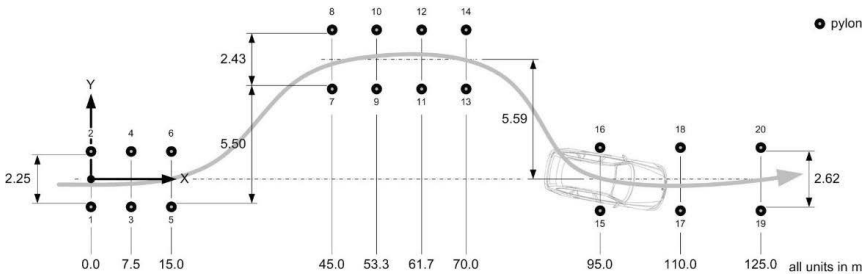


Figure 32. The double lane change maneuver with the dimensions and cones numbers (Oort, van, 2011).

The drivers were free to adapt their speed but were instructed to drive on the maximum velocity with which the double lane change could be driven without hitting any cones. The vehicle was kept in the third gear for all tests.

Immediately after each test, the driver filled in the assessment form (Appendix 1), to assess the subjective evaluation of the tire in that test. The aspects on this assessment form were set up in discussion with the professional tire test drivers. Table 3 contains the handling aspects that were graded by the drivers.

Table 3. Handling aspects for the subjective assessment of tires.

#	Aspect
1	Steering precision while cornering
2	Stability while cornering (no throttle change)
3	Stability while cornering (throttle change)
4	Yaw overshoot
5	Predictability
6	Yaw delay
7	Steering angle
8	Grip
9	Controllability
10	**Overall judgment**

The subjective assessments were performed with so-called blind testing, *i.e.,* the driver had no information of which tire was tested. All tires were covered with clothes, before, after and in between testing and were therefore not visible to the drivers. In accordance with the, at that time, used test practice of the professional test drivers, similar tires were assessed in groups, called batches. This resulted in three batches that included the following tires:

- Batch A: tires 1 and 2
- Batch B: tires 2, 3 and 4

- Batch C: tires 5 and 6

Tire 2 is included in two batches, batch B having the most similar tires, but because tire 1 cannot be tested alone in a batch, tire 2 was also added to batch A. The professional test drivers rated the handling aspects of the several tires in scores from 1 to 10, but these scores cannot be seen as absolute values. Although nowadays the professional test drivers rate tires in absolute scores from 1 to 10, at that time, this was not the case. Tires were compared on scores given by one driver within one batch and not between drivers and batches. Drivers could therefore use their own interpretation of a score and their own spread, making the ratings subjective (driver dependent). On the other hand, the scores are not entirely relative: the batch with summer tires was scored generally higher than the batch with winter tires. Comparison of absolute scores can therefore be done within batches for one driver, but comparison between batches is less reliable. Comparison of scores between drivers for the same tires in a bath can only be done on the level of best, intermediate or less good scoring tire. The subjective assessment scores for all assessments are given in Appendix 5. Each driver tested every tire. Both drivers performed a total of 7 assessments, where during one assessment 3 to 5 double lane changes were driven. During driving, the driver determined the number of double lane changes necessary for obtaining a good feel of the tire. In total, the drivers performed 60 double lane changes during testing that could be used for analysis. Fig. 33 gives an overview of the number of double lane changes divided in driver, batch and tire for this experiment.

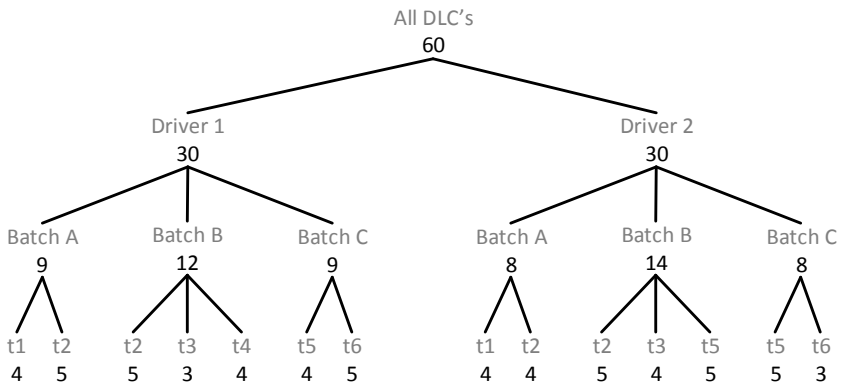


Figure 33. Total number of 60 double lane changes (DLC's) divided in driver, batch and tire (t).

4.7 Workload Experiments

The two field experiments that were set up to influence and measure driver's workload are described in this section: the slalom tests and the double lane

change tests. This section starts with describing the general methods used for these workload experiments on how the adapting driver theory was implemented, how the driver workload was influenced and how the driver's task performance and workload were measured.

4.7.1 Adapting Driver Theory Implemented

The adapting driver theory, as explained in section 2.4.4, is the basis for performing the experiments described in this section and therefore this theory applied to tire testing, is shortly summarized here. During tire handling assessment, the driver is assumed to act in the task related effort region *A*3 of the driver's mental workload curve (Fig. 19, page 33).

In this task related effort region *A*3, primary task demand measures are not sensitive for measuring task demand differences coming from different tires, because the performance does not change. Nevertheless, professional test drivers are able to feel and assess differences in this region. It is assumed that the driver adapts his behavior to different tires that produce different task demands to keep the task performance high. For example, the task demand is to drive a maneuver at high speed, without hitting any cones and to perform this with different handling tires. It is assumed the driver is forced to put in more effort, resulting in a higher mental workload, with a less handling tire than with a better handling tire to keep the performance high, *i.e.*, not to hit any cones. If the driver's mental workload could be measured and these measures are correlated to the driver's subjective assessment of the tires, this could give insight in this driver 'feeling' and could be used to give an additional measure of tire handling. To confirm these assumptions, two different experiments were performed, shortly referred to as the slalom tests, described in section 4.7.4 and the double lane change tests, described in section 4.7.5. In the next sections, the applied methods for influencing and measuring driver's workload are described.

4.7.2 Influencing Driver's Workload

Although during normal driving, next to the primary task of driving, other tasks are often present that can increase driver's mental workload, like talking on the phone or operating buttons on the media console, this is not the case during tire test driving. During tire test driving, the test driver focuses fully on the driving itself, to be able to compare different tires. Driver's mental workload for tire testing is therefore assumed to be related to the primary task of driving, which may be different for different tires. In line with this, for both experiments, the driver's mental workload was influenced by changing the primary task demand. Increasing the task demand in the task related effort region *A*3 also increases the mental workload of the driver. The independent variables chosen to manipulate this task demand, and therefore the driver's mental workload, were speed, tire and for the slalom experiment also the maneuver itself.

Speed Influence on Task Demand

Speed is chosen as a variable to change the task demand level. The higher the speed, the more demanding driving the maneuver is assumed to be. Fig. 34 shows two speed demands in the task related effort region, the low speed task demand where the performance is high (left hand black dot), combined with low driver mental workload (left hand grey dot) and the high speed task demand where the performance is still high (right hand black dot), but now combined with high driver mental workload (right hand grey dot). How the actual values for the speed task demands were derived, is explained in the sections describing the specific experiment. The labels 'low speed' and 'high speed' used in this research are relative speeds in the context of tire testing. The actual speed values for the low speed were between 80 and 100 km/h and for high speed between 90 and 110 km/h.

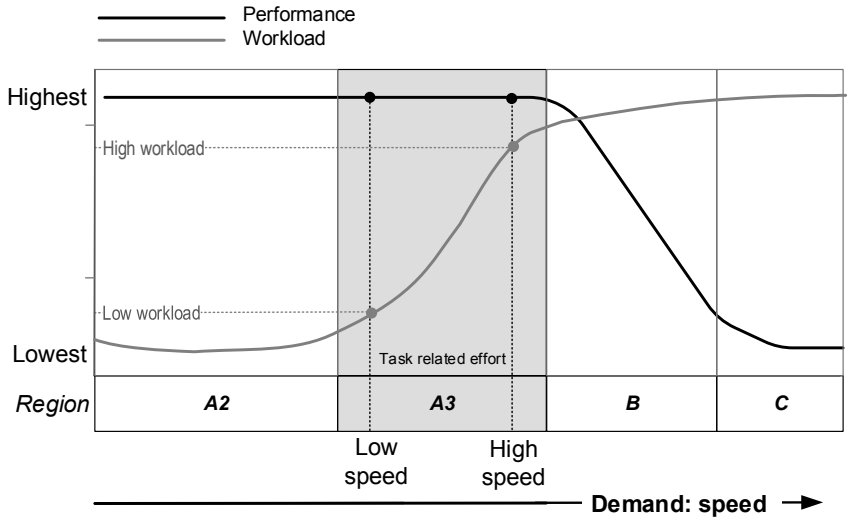


Figure 34. Low and high speed demands resulting in the same highest performance (black dots) but in low and high driver mental workload (gray dots).

For these two task demands, distinction cannot be made on the performance level, because the performance level is the same, both at the highest level (no cones hit) as can be seen for the black dots in the figure on the performance curve. There is however, a difference in mental workload, the grey dots on the mental workload curve. This implies that in the task demand region, distinction in demand level cannot be made based on performance differences, but can be made based on mental workload differences.

Tire Influence on Task Demand

A less good handling tire is assumed to impose a higher demand on the drivers compared to a better handling tire. Fig. 35 shows the expected performance curve for a less handling tire as a black, solid line. The performance for the high speed demand is high, hitting no cones. The grey solid line shows the

corresponding driver mental workload. For the high speed demand, the mental workload is high to keep the performance up (high workload less handling tire). The low speed demand also corresponds to high performance, but now combined with lower driver mental workload (low workload less handling tire). The same graphs are shown in the figure for a better handling tire with dotted lines. For this tire, the performance and mental workload curves and the corresponding regions (not shown in this figure, but shown in Fig. 34) all shift to the right compared to less handling tire. This results in less driver mental workload for this better handling tire for the same low and high speed demands compared to the mental workload of the less handling tire. Differences between these two tires cannot be derived from performance measures, while for both tires the performance level is high; differences can be derived from the driver mental workload: higher mental workload for the same speed demand indicates a less handling tire. Note that based on this graph the expected differences between driver's mental workload due to the different tires will be smaller than the expected differences in driver's mental workload due to different speed.

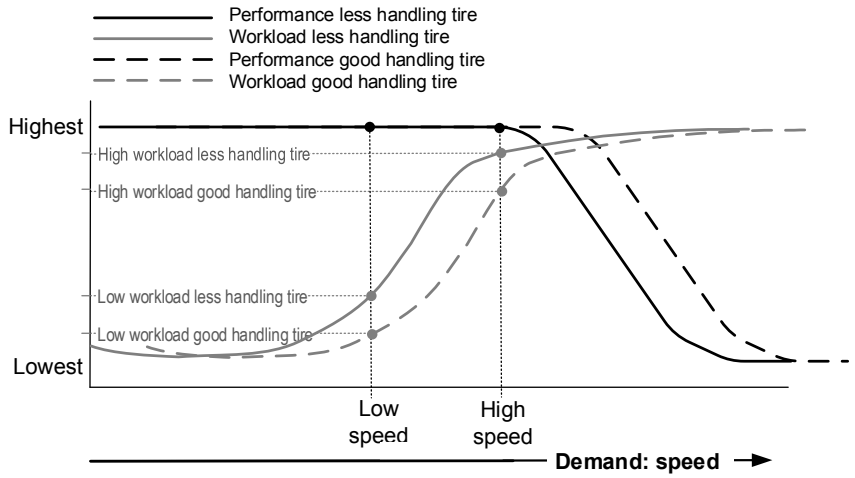


Figure 35. Low and high speed demand, both in region $A3$, resulting in different driver mental workload for different tires combined with equal high performance.

Increasing task demand causing increasing workload combined with constant high performance, is only valid in region $A3$, the task related effort region. If demand is increased, resulting in entering region B, the performance cannot be maintained on the high level, but will go down. This is the case when the handling of a tire would be worse compared to the less handling tire. In Fig. 36, the solid lines represent the performance and workload curves of the less handling tire, just as in Fig. 35. The same graphs are shown in the figure, but now for a worse handling tire, given with dotted lines. For this tire, the performance and mental workload curves and the corresponding regions (not shown in this figure, but shown in Fig. 34) all shift to the left compared to less

handling tire. For the low speed demand, the performance is high for both tires. However, this is accompanied with (considerable) higher driver mental workload for the worse handling tire (indicated in the figure with 'low' workload worse handling tire) compared that of the less handling tire. For the high speed demand, the workload for both tires is almost equal at the highest workload. Despite this high workload, the performance of the worse handling tire cannot be maintained on the highest level, indicating that, for this tire, the high speed task demand is in region B.

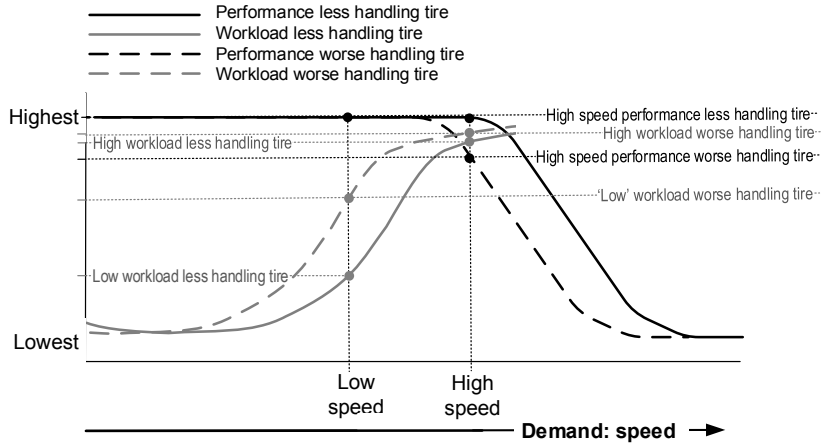


Figure 36. Low speed demand in region $A3$ resulting in different driver mental workload for different tires combined with equal high performance. High speed demand in region B, resulting in almost equal driver mental workload for different tires combined with different performance.

Based on Fig. 36, differences between the less and worse handling tires could be derived from performance measures. For the high speed demand, the performance of the worse handling tire is the lowest. As performance measures are in general easier to obtain compared to workload measures, this was also explored for this research. However, this was not a feasible option. Experiments with test drivers on a test track showed that it is not advisable to distinguish handling in this region B, while the high speed demand that corresponds to decrease of performance, despite drivers maximum effort, will make the vehicle easily become unstable with all the risks involved. In addition, the decrease of the performance happens very rapidly, making differences hard to measure, or as drivers stated: 'If it goes slightly wrong, it goes totally wrong!'

4.7.3 Measuring Driver's Task Performance and Workload

The driver's primary task performance level is measured by the number of cones that were hit during the maneuver, no cones hit was considered the highest performance.

The driver's mental workload was measured with measures from each of the three main categories mentioned in section 2.4.5, a subjective measure using a rating scale, two physiological measures based on heart rate measurements and a performance measure for the secondary task based on reaction time for the double lane change tests. In addition, the steering wheel measurements (section 4.4) were used to indicate mental and/or physical workload.

Rating Scale

For the subjective measure, the experienced workload was indicated by the drivers by using the Rating Scale Mental Effort RSME (Zijlstra & Doorn, 1985), see Appendix 2. This is a rating scale consisting of a 150 mm line marked with several anchor points, having a descriptive label (in the driver's mother tongue) indicating a degree of effort, *e.g.,* 2: absolutely no effort, 112: extreme effort. This rating scale is used for many studies for driver's mental workload (Cnossen, 2000; Hogema & Veltman, 2002; Waard, de, Kruizinga, & Brookhuis, 2008; Waard, de, 1996), because it showed good results, even compared to more complex methods (Verwey & Veltman, 1996). Although it was not assumed that the physical workload would differ significantly between tires, it could increase for higher speed, because a driver has to steer more rapidly at higher speed for the same maneuver. This change in physical workload could be mixed up with mental workload by the drivers. To let the drivers explicitly distinguish between these two, the drivers had to score mental and physical workload separately.

Heart Rate Measurements

As a physiological measure, an electro cardio graph (ECG) of the driver was made during driving the maneuver using skin electrodes. As explained in section 2.4.5, the IBI and HRV are measures derived from these heart rate measurements that are successfully used outside the laboratory for drivers. Fig. 37 shows an example of IBI data.

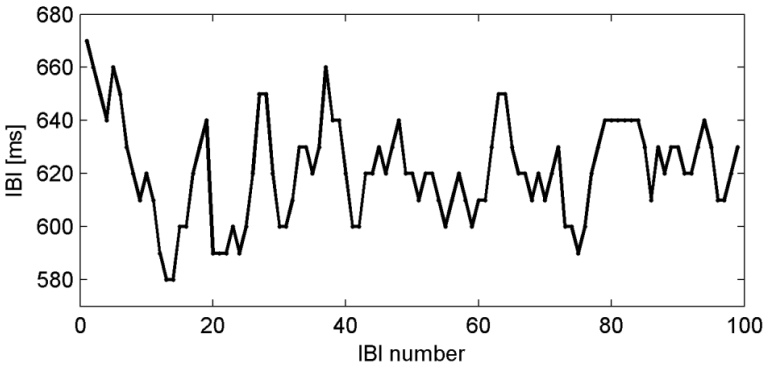


Figure 37. Example of Inter Beat Interval (IBI) data.

HRV is the variability of this IBI-data. As indicated in section 2.4.5, heart rate measures can be used if the duration of the workload is at least 30 s. Therefore, this measure could not be used during the double lane change tests, while

driving this maneuver takes only a few seconds. It is used in the slalom test, which length is specifically chosen, so that driving it would take more than 30 s

Secondary Task Reaction Time

Due to a primary task demand, part of the driver's mental workload will be allocated for keeping the primary task performance high in the task related effort region (Fig. 38). Depending on the level of the task demand, part of total available mental workload is available for performing a secondary task. The performance on this secondary task is assumed to decrease with decreasing mental workload available for this secondary task, which corresponds to increasing mental workload needed for the primary task. Measurement of the secondary task performance can then be used as indirect measurement of the driver's mental workload allocated for primary task performance. Decreasing performance on the secondary task indicates increasing driver's mental workload for the primary performance.

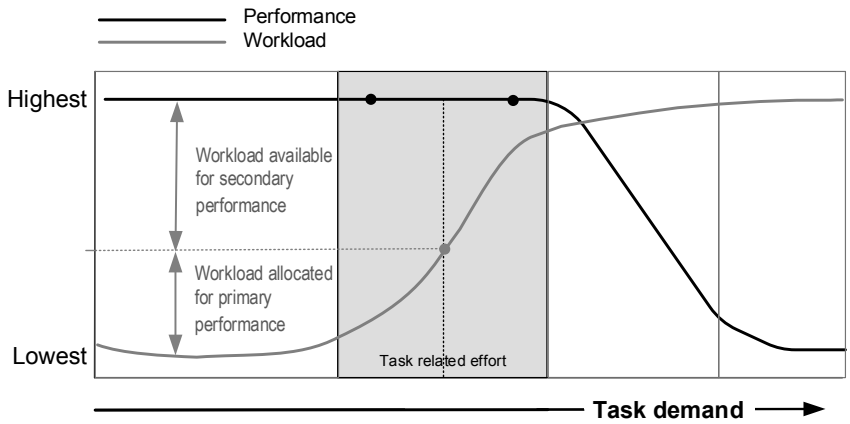


Figure 38. Workload allocated for primary performance and available for secondary performance.

A visual detection task is chosen as secondary task. Although the visual detection task is more sensitive to visual workload than to mental workload according to the multiple resource theory of Wickens (Wickens, 2008) it is also sensitive for mental workload. The peripheral detection task, where visual stimuli are presented to a driver with the task to react as fast as possible to these stimuli, the reaction time showed to be a good measure for secondary task to indicate mental workload, also for short durations of mental workload (Jahn, Oehme, Krems, & Gelau, 2005; Martens & Winsum, 2000; Verwey & Veltman, 1996). While driving the double lane change has a duration which is too short for heart rate measures, this secondary task is chosen for measuring mental workload for these tests. During the maneuver, the LED's lit up three times randomly in time during 0.4 s, in the field of view of the driver (Fig. 39).

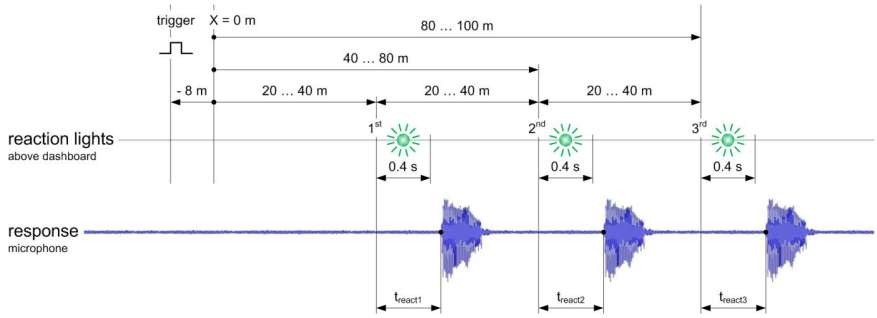


Figure 39. Scheme secondary task (Oort, van, 2012).

Although the peripheral detection task is meant to give visual stimuli in the peripheral field of the human view, studies with drivers showed that the position of the stimuli did not influence the measure (Martens & Winsum, 2000). The actual position of the LED's was chosen after testing with different drivers. Positioning of the LED's more to the side of the driver's view, resulted in missing almost all visual stimuli. As the reaction time was meant as secondary workload measure, missing stimuli by positioning of the LED's should be avoided. The drivers were instructed to react as fast as possible to these LED's with saying 'ja' ('yes' in Dutch), but with the explicit instruction to give always priority to the primary task, which was repeated several times to the drivers during testing to stress the importance.

This secondary task was implemented using a data logger, which runs separately from the main computer. A microphone, reaction lights (LED), optic cell and a small control box for the experimenter have been connected to the data logger. The optic cell on the rear bumper of the test vehicle was triggered with a reflector post positioned 10 m before the first maneuver cones, so the vehicle CoG was around 8 m before the first cone. This signal was also logged on the main computer and used for synchronization between main computer and data logger afterwards. Additionally, a video camera was installed that captured the reaction lights, steering wheel, and arms and head of the driver. Being a secondary task, it should have no effect on the primary task. To be able to verify if this secondary task had no confounding effect on the primary task performance, all (repeated) tests were done twice: with and without secondary task.

Steering Wheel Measurements

As indicated in section 2.4.3 steering measures derived from steering wheel measurements can be used to indicate both physical and mental workload.

4.7.4 Slalom Experiment

In this experiment, two nonprofessional, but highly skilled drivers participated. The task demand for the drivers was to drive a slalom between cones positioned in a circle having a radius of 100 m. This assured that the duration of the maneuver was at least 30 s, so heart rate measures could be used for indicating

mental workload. The maneuver was driven with a certain, predefined speed kept by the cruise control. To minimize the predictability of the maneuver and to avoid building up experience of the driver (which is both known to lower the workload) a curved path was used. This also prevented the driver to oversee the whole slalom and cone spacing at once, making it more difficult. For the same reason drivers were not allowed to watch each other's performance. By driving on cruise control, the driver was not able to alter the speed of the vehicle; hence, his only input was the steering wheel. After performing a test, test drivers switched, so the drivers had the same test circumstances and a certain time between the tests to prevent influences between successive tests. Drivers did not drink coffee, smoked or performed physical work between tests to prevent this influencing their heart rate.

To influence the dependent variable, the mental workload of the driver, the independent variable, the task demand, was changed by changing the difficulty of driving the slalom. By increasing the task demand for the driver until the point where he cannot keep up the high performance (no cones hit), the driver is expected to go through the task related effort region A3. In this region, the mental workload increases with increasing task performance to keep the high performance level. These assumptions are validated in section 5.1.

The task demand of driving the slalom was adjusted by changing the following independent variables:

- Speed: set by the cruise control to: 30, 45, 60, 65 and 70 (km/h). Increasing the speed is assumed to increase the task demand. The maneuvers driven at 30 km/h are also driven without cruise control, to compare workload with and without cruise control. Additional, a maneuver at maximum possible speed for the driver is driven without cruise control to determine the maximum demand.
- Cone spacing: 16 meters, 20 meters and (random, unpredictable for the driver) combinations of 16 and 20 meters. Smaller constant cone spacing and random cone spacing is assumed to give a higher task demand.
- Tire pressure: Tire 4 was used in this experiment with normal and adjusted tire pressure. For the adjusted case, the tire pressure at the rear is set from 2.7 to 4 bar to decrease the grip. Test driving showed that the vehicle behavior was more towards oversteer, which is assumed to give a higher task demand.

Each driver performed 12 maneuvers with varying speed, cone spacing and with normal and adjusted rear tire pressure to give different levels of task demand, resulting in 24 maneuvers. Table 4 gives an overview of the maneuvers.

Table 4. Slalom maneuvers driven by both drivers with normal (N) and adjusted (A) rear tire pressure.

Cone Spacing (m.) / Speed (km/h)	16	16/20	20
30 (no cruise control)	N		
30	N		
45	N		
60	N	N	
65	N,A		N
70		N, A	
Max (no cruise control)		N, A	

All three independent variables are assumed to also influence the physical workload. Higher speed would require faster steering for driving the same maneuver. Varying cone spacing would require more varying steering compared to constant cone spacing. This also holds for smaller constant cone spacing compared to larger constant cone spacing. With adjusted rear tire pressure, the vehicle behaves more towards oversteer, which makes the vehicle at higher speed less stable, requiring more corrective steering compared to understeer behavior. To let the drivers make explicit distinctions between mental and physical workload, both are rated separately.

4.7.5 Double Lane Change Experiment

In this experiment, two groups of drivers participated, one group of two professional tire test drivers and one group of six nonprofessional, but highly skilled drivers, but not familiar with tire testing.

The task demand for the drivers was to drive a double lane change (Fig. 40). The maneuver was driven with a predefined speed kept by the cruise control. The offset between the lanes was set to 5.5 m, instead of the 3.5 m prescribed in the ISO double lane change (ISO, 1999), to make the maneuver more challenging for the drivers without having to drive with very high speed. Unfortunately, due to the track humidity during the available test period for the professional drivers group, the lane change offset had to be set back to 3.5 m for safety reasons.

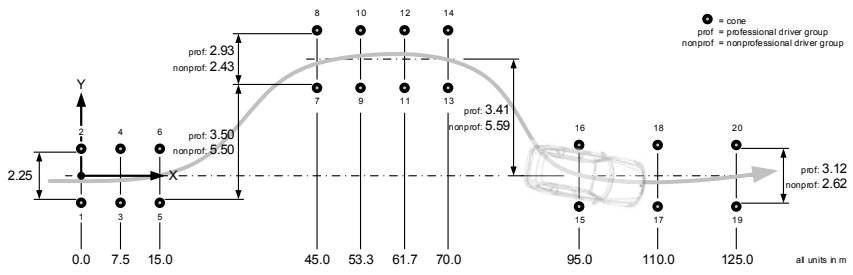


Figure 40. The double lane change maneuver with the dimensions and cones numbers (Oort, van, 2012).

To influence the task demand, the difficulty of performing the double lane change was adjusted by changing the independent variables speed (low and high speed) and tire (tire 2, 5 and 6). By adjusting these independent variables, the drivers are assumed to act in the task related effort region A3. All these assumptions are validated in section 5.3.

The tests were driven in test clusters, a test cluster being a combination of driver, tire and speed performing consecutive double lane changes. In every test cluster the maneuvers were repeated two (nonprofessional drivers group and afternoon of professional drivers group) to three (morning of professional drivers group) times and this was repeated with and without secondary task, giving 4-6 double lane changes in one test cluster. After each test cluster, the driver scored the experienced mental and physical workload. Each test cluster driven in the morning was also repeated in the afternoon. The order of testing a tire, driving with or without secondary task, the order of driving the speed, all these conditions were counterbalanced as much as possible, to reduce possible order effects. This resulted in data of 360 usable maneuvers, of which 120 were driven by the professional drivers and 240 by the nonprofessional drivers, evenly distributed over speed, tire and with/without secondary task.

Low and High Speed Values

As explained in section 2.4.4, the positioning of the performance and mental workload curves on the task demand axis is subjective. It depends on the difficulty of the task, which depends not only on the task complexity, but also on the driver. Therefore, the values for the low and high speed demands, given as setpoint to the cruise control, were determined separately for the drivers groups and for the morning/afternoon due to possible different circumstances. The speed values were determined so that the high speed value was the speed for which the drivers in that group were just able to drive the maneuver with the less handling tire 2, without hitting any cones. This assured that the high speed demand was near to the border between regions *A*3 and *B* (Fig. 34). A slightly higher speed resulted in decrease of performance, cones were hit, indicating entering region *B*. Drivers were therefore required to invest much effort for keeping the performance high (no cones hit), resulting in a high driver mental workload. The low speed value was then chosen in a way, so that driving the lane change without hitting any cones was less demanding for the drivers, but still challenging, requiring more mental workload than considered low, which would indicate the left border of region *A*3.

The actual values for high and low speed were chosen specific for both groups, because the groups differed in skill level and track. The change in track humidity between the morning and the afternoon of the professional drivers test day also caused a difference in the values for high and low speed. For all low speed values, the value was found 10 km/h lower than the corresponding high speed value. Table 5 shows the values of high and low speed for the two drivers groups.

Table 5. Low and high speed values driven during maneuvers.

Group	Maneuvers	Low speed (km/h)	High speed (km/h)
Professional drivers	First half	90	100
	Second half	100	110
Nonprofessional drivers	All	80	90

Tire Influence on Task Demand

From the subjective evaluations of the professional drivers (described in section 4.6) the average value for the overall judgment of handling evaluation of the three tires used in this experiment are given in Table 6. It should be noted that tire 2, a winter tire, was tested in a different batch than tires 5 and 6, both high performance summer tires, which implies that comparison of absolute scores is not completely correct, but it can give a good indication (for more information, see section 4.6). In general, high performance summer tires, being especially designed for good handling, can be expected to be better handling tires than winter tires.

Table 6. Average value of the overall handling judgment of the professional drivers group.

Tire	2	5	6
Overall Judgment Mean	6.75	7.25	7.75

Tire 2 as the less good handling tire, is assumed to impose a higher demand on the drivers compared to the summer tires. For 6 as the best handling tire, the imposed demand is expected to be lower. If the driver acts in the task related effort region *A*3 a higher demand is accompanied with a higher workload.

5. Validation of Assumptions Field Experiments

Several field experiments were performed for this research, as described in chapter 4. For all these experiments, certain assumptions were made. These assumptions are validated in this chapter for the subjective tire handling experiment, the slalom experiment and the double lane change experiment.

5.1 Validation of Assumptions Subjective Tire Handling Experiment

5.1.1 Speed and Performance Demand

Table 7 shows the mean speed driven during the first part (2 s before until 3 s after first cone) of the double lane changes for all tires. This shows that the double lane changes were driven with high speed, taken into account the large offset of 5.5 m. The increasing mean speed for batch A, B and C indicates that the drivers could drive the maneuver with higher speed for the better handling tires.

Table 7. Mean speed driven during first part of double lane change for all tires and mean value of the overall judgment score.

Batch	**A**		**B**			**C**	
Tire	**1**	**2**	**2**	**3**	**4**	**5**	**6**
Mean Speed (km/h)	81	85	85	86	86	92	92
Overall Judgment Mean	6.00	6.75	6.75	7.00	7.63	7.25	7.75

A histogram of the number of cones hit during driving the 81 double lane changes is shown in Fig. 41. This shows that almost all maneuvers were driven without hitting any cones or hitting only one cone. As explained in section 4.6, small, flat cones were used, which made their visibility for the driver less good, which explains why often a cone was hit without the driver noticing.

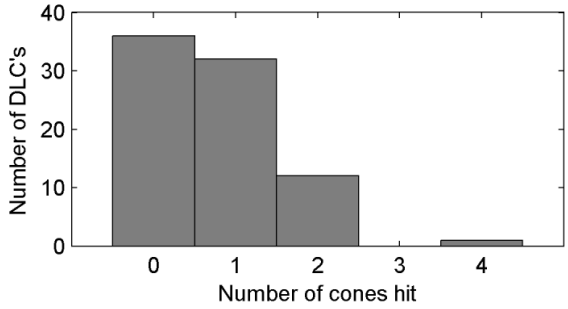


Figure 41. Histogram of the number of cones hit during driving the 81 double lane changes (DLC's).

5.1.2 Expected Tire Handling Performance

For this research, six tires were selected based on their expected handling performance as assumed by the tire manufacturer based on their data, test results and experience. Fig. 22 on page 43 shows the relative expected handling performance for all six tires, numbered from 1 to 6 related to the expected performance, 6 being the highest.

The following ISO open-loop test methods were performed:

- The steady state circular test (ISO, 2004) for analyzing steady state behavior.
- The step steer test (ISO, 2003) for lateral response analysis in the time domain.
- The pulse steer test (ISO, 2003) for lateral response analysis in the frequency domain.

The objective results (Monsma, Buning, Pauwelussen, & Bremmer, 2008) confirmed the expected handling performance of the tire manufacturer. In addition, experiments for subjective assessment of the tires, described in section 4.6, were performed. The mean value of the overall handling judgment of the two professional test drivers is presented in Table 7, showing that the score increases for increasing tire number, except for tires 4 and 5. This could be explained by the fact that the tires are tested in separate batches, where relative scoring was applied, as explained in section 4.6. As tire 2 was tested twice in two different batches, this could be taken as a reference tire. As this tire received the same mean score in both batches, this makes it justifiable to compare the absolute values of batch A and B. As no such reference tire is present in batch C, this cannot be done for this batch.

The results described in this section, validate the assumptions that the drivers were pushing the tires to the limits during driving the maneuvers, to be able to assess the tire handling behavior comparable to their regular testing.

5.2 Validation of Assumptions Slalom Experiment

5.2.1 Speed Demand

Table 8 shows the mean speed demand as set on the cruise control for the slalom tests and the mean actual speed driven.

Table 8. Mean speed demand and mean actual driven speed during slalom tests.

Mean Speed Demand (km/h)	30	45	60	65	70	Driver C max	Driver F max
Mean Actual Speed (km/h)	28.5	41.5	54.8	59	60.8	61.5	55.5

It can be validated that for increasing speed demand the actual driven speed was also increased. The actual driven speed was lower than the given setpoints to the cruise control, which for the lower speed is due to the bias of the cruise control. For the higher speed, especially for the speed demand of 70 km/h, this is probably also caused by severe cornering during the maneuver. This lowers the speed of the vehicle and is difficult to compensate for the cruise control.

5.2.2 Influence of Speed, Cone Spacing and Tire Pressure on Mental Workload

In the task related effort region *A*3, an increase in task demand is accompanied by an increase of mental workload. Therefore, in this region, increasing the task demand can be used to increase the driver's mental workload. The assumptions were made that increasing the speed, changing the cone spacing to varying values and/or changing the tire rear pressure from normal to adjusted would increase the driver's task demand and subsequently his mental workload. Fig. 42 shows three plots with results of the slalom maneuvers. On the *y*-axis the dependent variable, *i.e.,* the mental workload score for both drivers, is given. Each plot shows the dependency of the workload score on one of the independent variables, *i.e.*, speed, cone spacing and tire pressure. In plot a the mental workload score is plotted for both drivers for varying speed, for all tests with normal tire pressure and cone spacing of 16 m It shows that for both drivers the mental workload score increases with increasing speed. In plot b the mental workload score is plotted for both drivers for equal tests with different cone spacing: 16 m and varying 16/20 m. Lines connect values of equal tests, which are tests for the same driver with a speed of 60 km/h and normal tire pressure, so only the cone spacing is different for two connected values. This plot shows that the mental workload perceived by the driver is increasing when going from constant cone spacing to varying cone spacing. Not shown in a plot, because it only concerns one set of comparable tests, was the increased cone spacing from 16 to 20 meters for 65 km/h. As expected, the shorter cone spacing was scored with a higher workload. In plot c the mental workload score is plotted for both drivers for equal tests with different tire pressure: normal and adjusted. Lines connect values of tests for the same driver, the same speed (60 km/h, 70 km/h and max) and the same cone spacing (equal 16 m and varying 16/20 m) and so

only the tire pressure is different for two connected values. Note that adjusted tire pressure is only driven in tests with higher speed, so only these values could be compared. It shows that for all tests adjusting the rear tire pressure, causes a higher driver mental workload.

The deviating combination of driver F in this plot with the lowest mental workload score for normal tire pressure combined with the highest mental workload score for adjusted tire pressure is explained by the fact that during driving with normal tire pressure, the driver missed several cones halfway the slalom and as he stated on the form, could not re-enter, so gave up his effort and therefore scored a low value for mental workload.

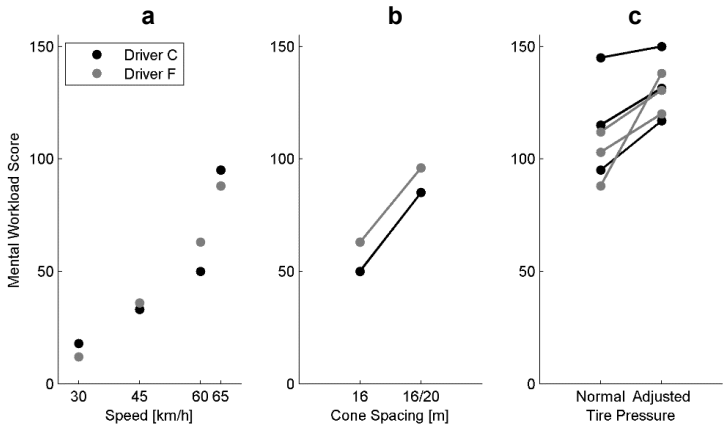


Figure 42. Mental workload scores for both drivers for:
a) Different speed values (all normal tire pressure and 16 m cone spacing).
b) Different cone spacing (all normal tire pressure and 60 km/h speed).
c) Different tire pressure (lines connect equal speed and cone spacing tests).

These results validate the assumptions that increasing the speed, changing the cone spacing to varying values or changing the tire rear pressure from normal to adjusted, increases the driver's task demand and subsequently his mental workload.

Workload Curve

Fig. 42a indicates that the relationship between speed and mental workload score is linear for 30 to 60 km/h, but after this speed, the workload increases at a much higher rate. This would indicate that for increasing speed demand, the workload curve in the task related effort region would follow this same curve. This is confirmed by the results of the double lane change experiment (see section 5.3.2)

5.2.3 Influence of Speed, Cone Spacing and Tire Pressure on Physical Workload

Next to indicating the experienced mental workload, the drivers were also asked to indicate the experienced physical workload, mainly to make sure that they

made a conscious distinction between these two. Fig. 43 shows the physical workload scores for both drivers for speed, cone spacing and tire pressure for the same tests as shown for mental workload scores.

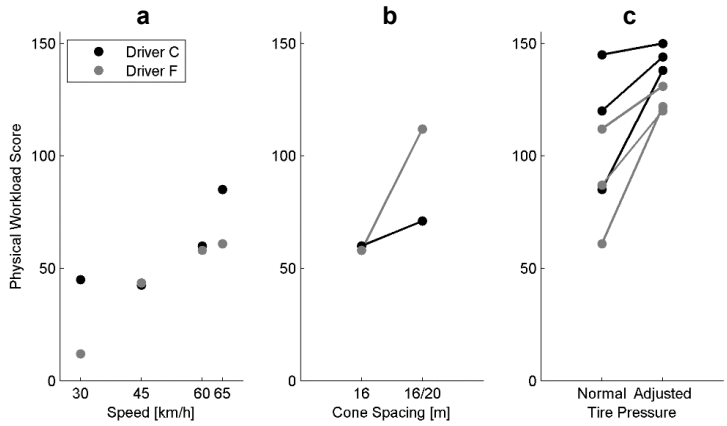


Figure 43. Physical workload scores for both drivers for
a) Different speed values (all normal tire pressure and 16 m cone spacing).
b) Different cone spacing (all normal tire pressure and 60 km/h speed).
c) Different tire pressure (lines connect equal speed and cone spacing tests).

These results show that increasing the task demand also increases the physical workload.

5.2.4 Task Related Effort Region A3

Fig. 42 shows that the slaloms were driven from low mental workload (Fig. 42 a low speed values) to highest mental workload (Fig. 42 c adjusted tire pressure), which corresponds to the task related effort region A3. The speed at which the task performance decreased, by hitting one or more cones, for driver C was around 65/70 km/h on cruise control and for driver F this was 60/65 km/h. These values correspond to the values of the average speed (after correcting for the cruise control bias) that the drivers had during driving the slalom without cruise control. These results validate the assumption that the drivers acted in the task related effort region *A*3 during driving the maneuvers.

5.3 Validation of Assumptions Double Lane Change Experiment

The results of the double lane change experiment are presented in graphs were the mean value of the dependent variable of interest, *e.g.*, the driven speed or the workload score, is plotted on the *y*-axis. This mean value is averaged over all the independent variables of interest, *e.g.*, tires or drivers group. For example, when comparing a certain workload score for different tires for both drivers groups, the mean value is averaged over all maneuvers driven with each of the three tires of all professional drivers and over each of the three tires of all

nonprofessional drivers, giving six mean values for this driver parameter, grouped by drivers group and tire. Speed and tire are ratio variables, as tire represents the handling qualification as is taken from the subjective mean score for the overall handling judgment (Table 6), therefore the mean values are connected with a line to show the trend. The mean values are plotted with an error bar representing the standard error of the mean (SEM) (Field, 2009), which indicates how accurate the estimation of the mean is likely to be. The calculation for the SEM for a group of N values x_i is :

$$\text{SEM} = \frac{s}{\sqrt{N}} \tag{11}$$

with s the (sample) standard deviation:

$$s = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}} \tag{12}$$

with $\bar{x}$ the mean value. This shows that SEM gets smaller as N gets bigger, so the estimate of the mean becomes better when N gets larger. Note the following two effects this has on the graphs of the results. First, the number of the maneuvers driven by the nonprofessional drivers group (240) is twice as large as the professional drivers group (120), so with comparable standard deviation in a group of values the SEM error bars will be larger with the professional drivers group than with the nonprofessional drivers group. Second, the more the results are grouped, *e.g.,* by drivers group, tire and speed, the smaller N becomes for a group, the larger the SEM error bars are likely to be. Both effects are correct: less values in a group results in a less good estimate of the mean.

5.3.1 Low and High Speed Demand

The high and low speed values for the two drivers groups are given in Table 5, page 62. These values were determined at the test track and served as the setpoint for the cruise control. Fig. 44 shows the mean values of the measured speed, grouped by drivers group, tire, speed demand and test cluster.

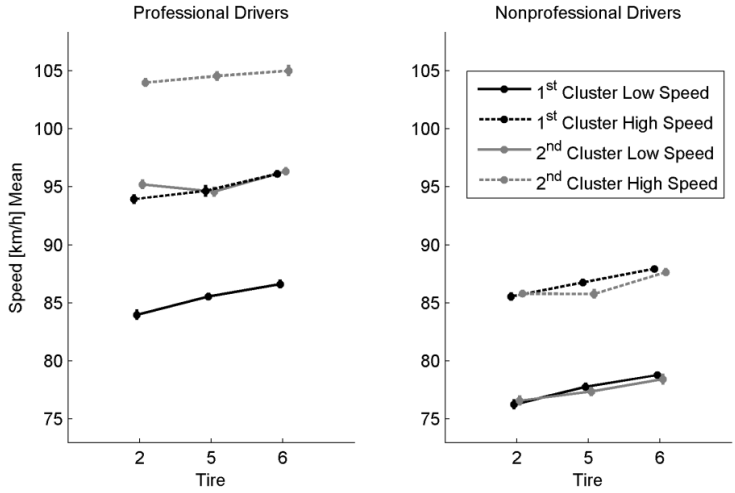


Figure 44. Mean values of the measured speed values, grouped by drivers group, tire, speed and cluster.

This figure shows that the actual driven speeds were around 5 km/h lower than the given setpoints to the cruise control, which is due to the bias of the cruise control. This does not influence the results, because for this research the (relative) speed demand connected to the position in the task related effort region *A*3 is important, not the absolute speed value. There is also a small tire dependency in the speed, which can be also related to the cruise control, where the actual speed is derived from the wheel rotational speed. The same rotational speed for tires having different effective rolling radii will result in the same speed for the cruise control, but different actual speeds. The increase of effective rolling radius of tire 2, 5 and 6 is respectively around 1 and 2 percent, which explains the same increase in percentage in the speed.

5.3.2 Influence of Speed and Tires on Driver's Mental Workload

The independent variables speed and tire are chosen to influence the level of the driver's task demand. In the task related effort region *A*3 (Fig. 19), an increase in task demand is accompanied by an increase of mental workload. Therefore, the assumption was made that in this region increasing the speed or changing the tire from 6 to 5 to 2 increases the driver's task demand and subsequently his mental workload. Fig. 45 shows the results for the mean value of the mental workload score, grouped by drivers group, tire and speed.

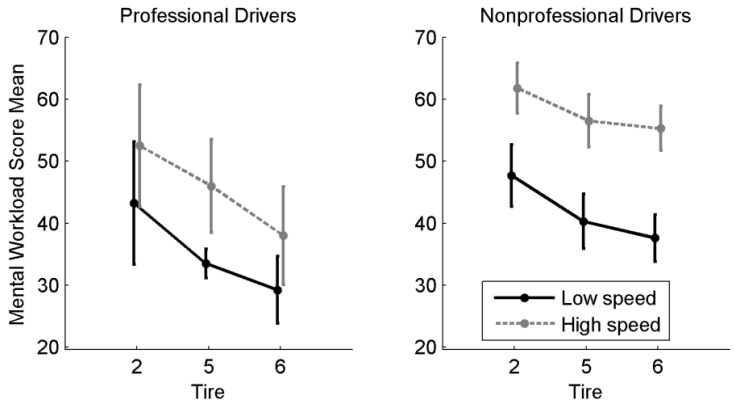


Figure 45. Mean values of mental workload scores, grouped by drivers group, tire and speed.

This figure shows that the mean mental workload scores are higher for high speed than for low speed for both drivers groups and all tires. This influence seems stronger for the nonprofessional drivers, having no overlap for the SEM error bars, than for the professional drivers, having overlap for tires 2 and 6 for the SEM error bars, but can also be owing to the group size differences. The higher absolute values for the nonprofessional drivers compared to the professional drivers, could be explained by the fact that for the professional drivers, this high speed tire testing is almost a 'daily routine', whereas for the nonprofessional drivers this testing was new and challenging, especially for the high speed demand.

The mental workload decreases from tire 2 to 5 to 6 for both drivers groups, with for the nonprofessional drivers and professional drivers for low speed, the difference between tire 5 and 6 smaller than between 2 and 5. This could be related to the fact that tires 5 and 6 are both summer tires and tire 2 a winter tire. These results validate the assumption that increasing the speed or changing the tire from 6 to 5 to 2 increases the driver's task demand and subsequently his mental workload.

The differences between the mental workload score for different speed are assumed to be larger than for different tires (Fig. 35). This is the case for the nonprofessional drivers, but not for the professional drivers, where these differences are similar. This can be explained by the fact that the professional drivers, being very experienced in tire testing, are better able to feel relative differences between tires than the nonprofessional drivers, having almost no experience in tire testing.

5.3.3 Task Related Effort Region A3

In section 2.4.4 it is explained that the assumption is made that during the performed tests, the test driver acts in the task related effort region *A*3 of the performance and workload curve as function of demand (Fig. 19). In this region, increasing task demand shows constant (high) performance of the driver, but at

the costs of increasing driver mental workload. In the previous text, the increasing mental workload due to increasing task demand (by changing the speed and tires) is shown, which corresponds to driving in the task related effort region A3, as this is the only region for this to be valid. In addition, in this region with increasing task demand, the task performance should also stay on the same high level, where a small decrease is indicating the border between region *A*3 and region *B*. This is shown in Fig. 46 by comparing the performance of the drivers, being the number of cones hit, during increased task demand, *i.e.,* increasing tire number. Maximum performance is no cones hit during driving the maneuvers.

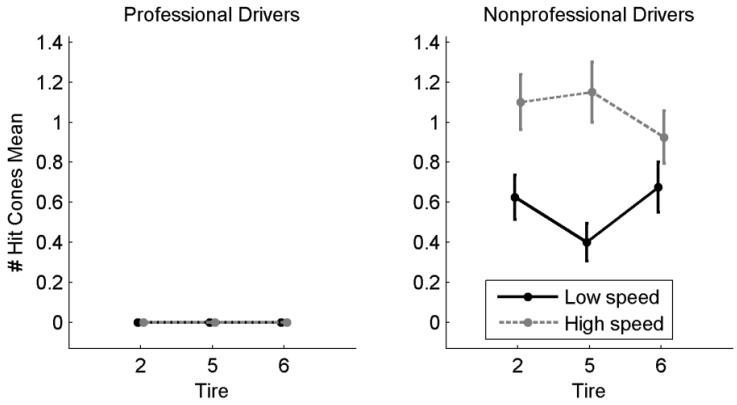


Figure 46. Mean values of the number of cones hit, grouped by drivers group, tire and speed.

This figure shows that the performance for the professional drivers stayed on a constant, highest level: no cones hit during any maneuver. For the nonprofessional drivers the high performance at low speed demand can be seen to be around the mean value of 0.4 – 0.6 cones hit during a maneuver, increasing slightly to 0.9 – 1.1 for the high speed demand, indicating entering the region B. Lowering the high speed value to shift the task demand a little more to the left, moving more into region A3, was not feasible during testing, as the drivers indicated during test driving this speed, that this considerably lowered their workload, which had the risk of not being near the border of *A*3 anymore. These results validate the assumption that the drivers acted in the task related effort region *A*3 during driving the maneuvers.

Workload Curve

Remarkable in Fig. 45 are the relative low absolute values of the mental workload score for the high speed demand, considering the full scale for the workload scores runs from 0 to 150. As explained previously, the speed values for high speed were chosen so that this driver group was near the border of the task related effort region *A*3 and the overload region *B* with maximum mental workload combined with primary performance decrease. Therefore, for high speed, high workload scores were expected, corresponding to Fig. 34 and

corresponding with the results found for the workload scores of the slalom experiment. Discussing these results afterwards with the drivers, revealed that the drivers experienced a huge increase in workload with a small increase of the speed above the high speed demand. This is consistent with the slalom results (Fig. 42a), where for speed demands above 60 km/h the mental workload increased at a much higher rate than for speed demands below 60 km/h. Comparing the change in workload for the tires in Fig. 45, shows that for the low and high speed the same trend is present. This indicates that the impact of changing the tire in terms of workload is the same for the low and high speed, which was as expected (see Fig. 35). This indicates linearity of the workload curve between the low and high speed demand. This is also consistent with the slalom results for the speed demands below 60 km/h. Combining these results suggests the workload curve in the task related effort region *A3* to follow the solid line in Fig. 47 instead of the dotted line, which was assumed (Fig. 34, page 53). The curve increases linearly but only until a certain workload (for the experiments, this is around a workload score between 60 and 70), after this value, the increase of workload with speed becomes very large, going to the maximum workload with a small speed increase.

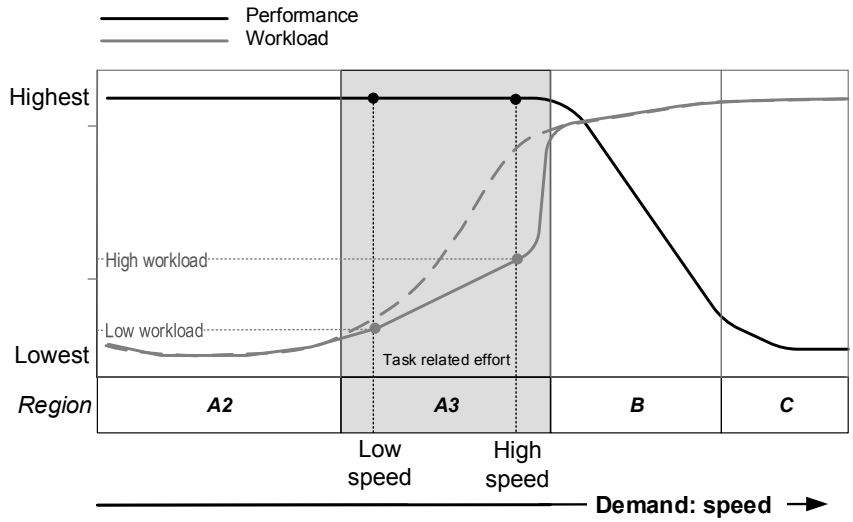


Figure 47. Adjusted workload curve (solid line) in the task related effort region compared to the original workload curve (dotted line).

5.3.4 Influence of Speed and Tires on Driver's Physical Workload

Next to indicating the experienced mental workload, the drivers were also asked to indicate the experienced physical workload, mainly to make sure that they made a conscious distinction between these two. Fig. 48 shows the results for

the mean value of the physical workload score, grouped by drivers group, tire and speed.

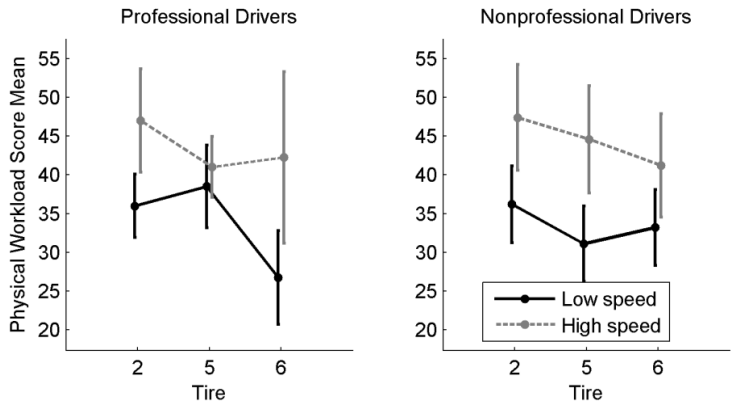


Figure 48. Mean values of the physical workload scores, grouped by drivers group, tire and speed.

This figure shows that the mean physical workload scores are higher for high speed than for low speed for both drivers groups and all tires, but the distinction is less clear than for mental workload. The tire dependency is not clear; the mean physical workload scores go down for the nonprofessional drivers but only for high speed. In addition, the length of the error bars compared to the mental workload graph (same value for N) indicates more deviation in scores. These results validate the assumption that speed has some influence on physical workload, but tires have almost no influence.

5.3.5 Secondary Task Influence on Primary task

During the experiment, tests were carried out with and without a secondary task, with the objective to see if driver mental workload is correlated with the secondary task reaction time of the driver. This secondary task should not influence the primary task of the driver, *i.e.,* driving the double lane change without hitting any cones. To verify if there is no confounding effect of this secondary task on the primary task, the primary performance measure - the number of cones hit during the maneuver - is compared for maneuvers driven with and without a secondary task. This is shown in Fig. 49, where the mean value of the number of cones that were hit during driving the double lane change are presented for all maneuvers driven without and with secondary task, grouped drivers group.

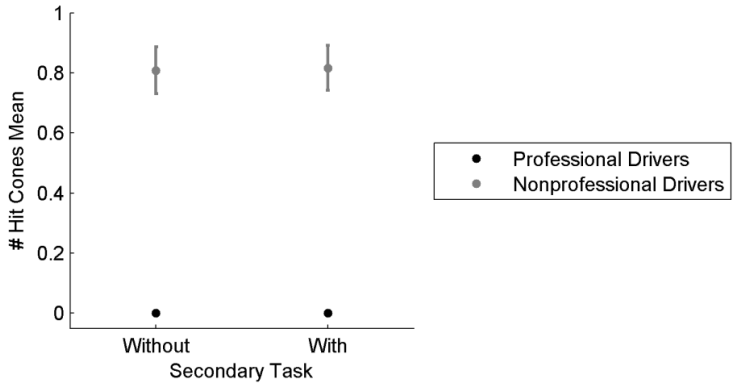


Figure 49. Mean values of the number of hit cones during driving the maneuver with and without secondary task, grouped by speed and drivers group.

The figure shows that there is no difference in the primary task performance for driving without and with a secondary task for the professional drivers, no cones hit for all maneuvers. The nonprofessional drivers also has similar results for the primary task performance for driving without and with secondary task, the mean number of cones hit being both around 0.8 and with almost identical SEM. This validates the assumption that adding a secondary task did not influence the primary task performance. Therefore, all maneuvers, regardless if driven without or with secondary task, are taken into account for estimating the driver model parameters.

5.4 Conclusions

All, except one, assumptions made for the experiments performed in this research are shown to be valid. The only exception is the workload curve in the task related effort region. The results indicate that this curve is not increasing linearly over the whole range, but only until a certain workload (for the experiments, this is around a workload score between 60 and 70), after this value, the increase of workload with speed becomes very large, going to the maximum workload with a small speed increase. This adjusted workload curve is shown in Fig. 47.

6. Driver's Handling Assessment using a General Regression Neural Network

This chapter describes the first method taken in this research for understanding a driver in vehicle handling as explained in section 3.1: predicting driver's subjective handling assessment based on vehicle dynamics measurements with a General Regression Neural Network (GRNN). More specific, a GRNN is developed for prediction of driver's subjective evaluation of tire handling on a test track (output), based on objective measures taken during a double lane change maneuver (input).

For a more detailed introduction to relating objective to subjective measures, the reader is referred to section 3.1; here, only a short summary is given. The process of driver's subjective assessment of the vehicle handling behavior can be seen from a system modeling view. The vehicle handling behavior is regarded as input to the driver. He perceives this input and assesses this handling behavior, which results in his subjective evaluation, the output. If this relationship between input and output could be modelled sufficiently, this can:

- Predict subjective evaluation based on (real or simulated) vehicle handling behavior.
- Give additional numerical output for supporting and comparing subjective assessments.
- Give information about what vehicle behavior is important for the driver in his handling assessment, by analyzing the related measures.

Therefore, the input is converted to characteristic data, the objective data, and the output consists of the driver's scores for several handling aspects (see Table 3, page 50), the subjective data.

The GRNN developed for the prediction, is a special kind of Artificial Neural Network (ANN). Fig. 50 gives an overview of the steps taken for this method, which will be described in this chapter. The shaded area shows the specific implementation for the method described in this chapter of the general model given in Figure 20 of section 3.1.

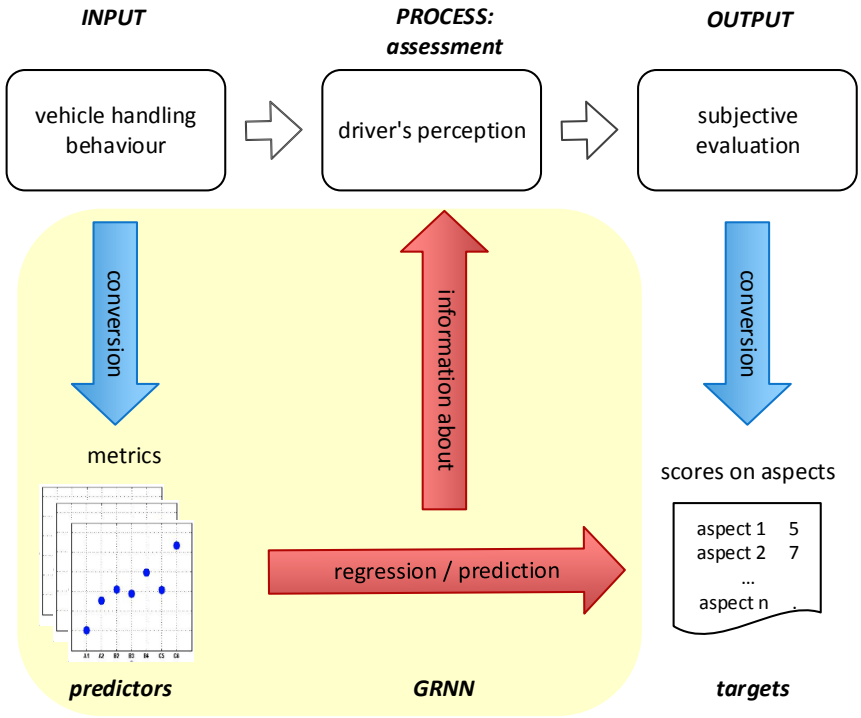


Figure 50. A General Regression Neural Network for prediction of aspect scores based on metrics.

Input and output data for the GRNN are described in section 6.1. Section 6.1.1 describes how the vehicle handling behavior is converted to characteristic data, here referred to as 'metrics', which can be used as inputs for the GRNN, the predictors. The outputs of the GRNN, the targets, are described in section 6.1.2. Section 6.2 describes the use of regression analysis for predicting handling assessment, what limitations are present and explaining the choice for using a GRNN as regression method. The GRNN itself is explained in general in section 6.3, where the specific implementation for this research is described in section 6.4. Results and discussion are given in section 6.5 and this chapter ends with conclusions in section 6.6.

To develop a regression model for prediction of driver's subjective evaluation scores, data are used from the experiment described in detail in section 4.6 and briefly summarized here. In this experiment, two professional test drivers performed subjective assessments of tire handling for six different tire sets (shortly referred to here as 'tires'), divided in three batches, A, B and C. Within each batch, tires were compared. In batch A, winter tires 1 and 2 were assessed, in batch B winter tires 2, 3 and all season tire 4 and in batch C summer tires 5 and 6. Each assessment consisted of driving several double lane changes and circles of 100-meter radius, the latter with and without throttle changes, the throttle controlled by the driver to his own convenience. After each assessment, the driver rated 10 handling aspects including an overall rating.

6.1 Input and Output Data

A regression model predicts one or several outcome variables, the targets, from one or several input variables, the predictors. To convert the input - the vehicle handling behavior - to predictors that can be used in a regression, the vehicle dynamics measurements taken during the handling assessments are converted to characteristic data, the 'metrics' (left-hand side of Fig. 50). This conversion is described in section 6.1.1 and resulted in 29 metrics. The output - the subjective evaluation - is quantified by the driver to scores on 10 handling aspects (right-hand side of Fig. 50), described in more detail in section 6.1.2. These metrics and scores, or a subset of both, can be used as respectively predictors and targets in a regression. Because there are six different tires, where tire 2 was driven in two separate batches, there are seven assessments for each driver. Although each driver assessed tire 2 twice, these two assessments are not duplicates. Tire 2 was driven in different batches on separate days, which could give differences in behavior and evaluation, resulting in different tire assessments. Each assessment resulted in a specific combination of 29 metric values and 10 aspect scores. In the context of regression, one assessment giving a specific combination of 29 metric values and 10 aspect scores is also referred to as one 'observation'. This gives seven assessments or observations for every driver, each having its specific 29 metric values and 10 aspect scores. These seven assessments are identified by the combination of the batch letter and tire number: A1, A2, B2, B3, B4, C5 and C6.

Next to these two data sets, representing the seven assessments of the two drivers, an additional data set was derived, to represent the mean of the two drivers, referred to as 'driver mean'. A possibility for deriving this driver mean data set could be to include all observations of both drivers in one data set, with the advantage of having more observations available for regression for the driver mean, which in general enhances the regression. However, the problem with this approach is that these observations do not represent similar driver assessment processes. Due to drivers' differences in driving style, (spread of) scoring and perceptions, the observations mutually differ more than for one driver alone. In this case, the underlying regression function represents both drivers' subjective assessment processes together in one data set. This results in a (much) less smooth underlying regression function, which - next to making regression more difficult to perform – does not represent what we actually want, *i.e.,* the mean of the drivers' subjective assessments. Therefore, another approach is taken; the driver mean data set was calculated by taking the mean values of the predictors and targets for both drivers. This resulted in an additional driver mean data set of also seven assessments.

To derive the metrics, the vehicle dynamics data of the double lane changes are used. This maneuver is especially useful for subjective evaluation of cornering behavior (ISO, 1999). Having well-defined boundaries, this maneuver can be driven consistently making the data very well comparable, in contrast to the driven circles, where the drivers were allowed more freedom to use the throttle and choose their path. Handling aspects 2 and 3, respectively stability while cornering with no throttle change and with throttle change, were

specifically determined while driving the circles. This behavior is not included in the metrics derived from the double lane changes, but will be predicted by the GRNN. Therefore, the assumption is that these aspects will not be predicted as well as the aspects related to driving the double lane change. Remarkably, the results invalidated this assumption; the GRNN was able to predict these aspects well from only double lane change metrics (see section 6.5.1).

6.1.1 Predictors

The conversion from vehicle handling behavior to metrics, which could be used as predictors, is explained in this section. As described in section 4.4 the test vehicle was equipped with instruments to measure various vehicle dynamics signals during driving. To capture the total dynamic behavior of driving a double lane change, the measured signals were taken from two seconds before the start line of the double lane change (position of the first cones marking the double lane change) to seven seconds after. For every signal, these nine seconds of the maneuvers were plotted to see if distinctive behavior between tire assessments was present. Various, often overlapping, parts were defined for every signal. The definition of these parts was an iterative process that is explained next, but as a first step, this was based on the most distinctive signal behavior for different tire assessments in this part. Fig. 51 shows an example, where two time series of the steering wheel angle are plotted, belonging to double lane changes driven by the same driver for two different tires. For this signal, there are five parts defined given in Table 9. For this signal, the parts capture mainly the several peaks in the signal. As can be seen, these peaks are different for the two tires. In this case, they are more pronounced for tire 1, driven in assessment A1, compared to tire 6, driven in assessment C6, for driving the double lane change. In addition, because the signal belonging to tire 1 is lagging behind compared to the signal of tire 6, the driven velocity for tire 1 compared to tire 6 was lower. This means that the driver drove and steered differently for these tires.

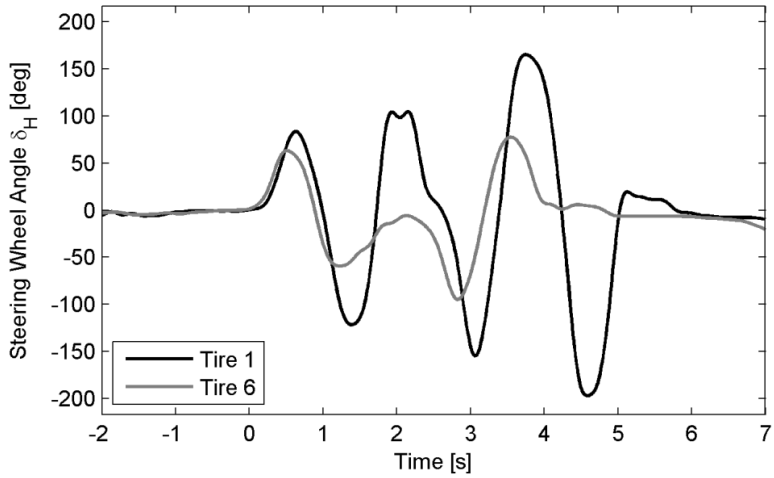


Figure 51. Two steering wheel angle signals belonging to double lane changes driven by the same driver with different tires.

Table 9. Signal parts of steering wheel angle.

Signal Part Steering Wheel Angle	Time t [s]
I	0.0 - 1.5
II	1.0 – 2.0
III	1.5 – 3.0
IV	2.0 – 4.0
V	3.0 – 4.5

After defining these signal parts as a first step, the following four calculations were done for each part: the maximum value, the minimum value, the integral for the positive values and the integral for the negative values. The maximum and minimum values are directly related to the signal peak values. The integral calculations are done to capture (partly) the speed of change of a signal, without using the noise-sensitive derivative of a signal. For example, even when signals in a part have similar peak values, they could have different behavior regarding the speed at which this peak is reached resulting in a narrow or wide peak. The latter will result in a larger integral value than the first.

In each assessment (one batch-driver-tire combination) several double lane changes were driven. The mean values of the calculations for these driven double lane changes in one assessment were taken to serve as a 'candidate metric'. By comparing the signal parts for the tires for a driver and plotting the corresponding candidate metrics (same signal, same signal part and same calculation), the suitability to be used as a metric was assessed. If the signal parts for the tires were overlapping much, resulting in the candidate metrics having similar values for different tires, this calculation was not taken as a metric, as it could not provide information to distinguish between different tires. If the signal parts were distinctive for the tires, for example, different peak

values, resulting in the candidate metrics having different values, this calculation was seen as a useful metric, as it could capture the distinguishing behavior between tires.

For all candidate metrics, the parts have been adjusted to optimize the metric to obtain the most distinguishing values of this metric for different tires. This iterative process resulted in 29 metrics for seven signals, *i.e.,* lateral acceleration, vehicle slip angle, yaw angle, yaw rate, steering wheel angle, steering wheel torque, lateral velocity and longitudinal velocity. For all signals, tables with the defined signal parts together with a plot of the signals for two different tires are given in Appendix 3.

All metrics were calculated for all tire assessments for both drivers and in addition, the mean value of both drivers for every assessment was taken. To eliminate the negative effects of different physical quantities used for the metrics and different scales of magnitude of the metric values, for each metric the values are standardized to zero mean and unity standard deviation. This gives a metric dataset of three dimensions: 3 x 7 x 29, for every driver (driver 1, driver 2 and driver mean), for every tire assessment (A1 to C6), 29 metric values were calculated that could serve as predictors in a regression. Appendix 4 contains scatter plots of the metrics and of the standardized metrics.

6.1.2 Targets

The conversion from subjective evaluation to handling aspects scores is done by the driver immediately after each assessment, by giving scores for 10 handling aspects. See Appendix 5 for an overview of all scores. Next to the scores of both drivers, the mean value of the drivers' scores was taken for every aspect of every driver. This gives a data set with three ‘drivers’, driver 1, driver 2 and driver mean, with each driver having seven tire assessments with 10 aspect scores that could serve as one or more targets in a regression.

Relative scores used as absolute scores

Although nowadays the professional test drivers are rating the aspects in absolute scores from 1 to 10, at that time, they did not rate in absolute scores, although the ratings were also from 1 to 10. Tires were compared for one driver within one batch and not between drivers and batches, making the ratings absolute within one batch of one driver, but relative between batches and drivers. Drivers could also use their own interpretation of a score and their own range of used scores, making the ratings more subjective (driver dependent). On the other hand, the scores were also not entirely relative: the summer tires were scored higher than the winter tires, but not as much as could be expected if all tires in all batches would be rated in absolute scores. This was also confirmed by the test drivers. Table 10 shows the mean score of all aspects for every driver for every assessment. Assessment B4 is scored higher than C5 by both drivers, for driver 2 it also scored higher than C6.

Table 10. Mean score of all aspects for every driver for every assessment.

Assessment	A1	A2	B2	B3	B4	C5	C6
Mean score driver 1	6,30	6,80	6,90	7,23	7,50	7,45	7,70
Mean score driver 2	6.33	6.65	6.70	6.70	7.25	6.85	7.10
Mean score driver mean	6,31	6,73	6,80	6,96	7,38	7,15	7,40

For a regression model, it is important to have a large data set available to obtain a reliable regression model. If the ratings were treated as relative ratings, so a regression could only be performed within one batch, this would reduce the data set too much, having only two or three tire assessments per driver per batch. Therefore, the ratings are treated as absolute ratings, although this is not entirely justified. Driver dependence of the ratings is largely removed by standardizing the scores for each aspect for every driver to zero mean and unity standard deviation. After standardization, the ratings have been used as if they were absolute ratings for this ANN, resulting in a target data set of three dimensions 3 x 7 x 10, for every driver (driver 1, driver 2 and driver mean), for every tire assessment (A1 to C6), there are 10 aspect scores that could serve as targets in a regression.

6.2 Regression Analysis for Predicting Handling Assessment

Regression analysis studies the relationship between predictors and targets with a regression model where the model parameters are estimated from the data, consisting of several observations (one observation being one combination of predictors and targets). The first step is choosing an appropriate regression model. This model is chosen according to an assumed relationship or behavior of the variables. When using experimental data, these data have a certain amount of uncertainty in it, like measurement noise or other random influences, which make the data less precise. Therefore, the regression analysis objective is to capture the underlying relationship or behavior of the data with the model, not to fit the model exactly on the data, as interpolation does. When the regression model is chosen, the second step is to fit the model to the data by estimating the values of the model parameters. For our experiment, we want to fit such a model to the experimental data, which can have up to 29 predictors and 10 targets, resulting in a high dimensional space. One observation or tire assessment, being a combination of values for the predictors and targets, can be seen as one data point in such a high dimensional space.

6.2.1 Limitations Regression Analysis

There are some limitations when using regression analysis for prediction.

Choosing Inappropriate Model
In general, the choice for the regression model is an important first step. If the chosen model is not appropriate for representing the underlying behavior, even the best fit will result in poor predictions. Hence, if little is known about the

underlying behavior, there is a risk of ending up with an inappropriate model due to improper model choice.

Poor Prediction Outside Data Range
Even with a proper model and a good fit to the data, the predicted behavior according to the regression model outside the data range used for fitting the model can become very poor. Especially for higher order models, the curve or hypersurface can show extreme behavior. Consequently, prediction of targets should only be done with predictor values that lie inside the data range used for the model fitting.

Poor Prediction of Higher Order Models between Data Points
Another disadvantage of fitting especially higher order models to data is the possibility of poor prediction between data point. Fitting only lower order models is not always a solution, if the assumption of such a lower order model is not justified or the data points are scattered in a way that the fit becomes very poor, often using a higher order model cannot be avoided.

6.2.2 Limitations Handling Assessments

For a regression with several predictor variables and several target variables, the general multiple multivariate regression model is given by:

$$\boldsymbol{t}_i = f(\boldsymbol{x}_i, \boldsymbol{\theta}) + \varepsilon_i \tag{13}$$

with

$\boldsymbol{t}_i$	the vector of target variables for observation i
f	the (non)linear regression function
$\boldsymbol{x}_i$	the vector of predictor variables for observation i
$\boldsymbol{\theta}$	the scalar or vector of unknown parameters
ε_i	the error between the calculated model outcome variable and the target variable of observation i
$(\boldsymbol{x}_i, \boldsymbol{t}_i)$	data point, combination of predictors and targets for observation i.

This model contains one or more unknown parameters in the vector $\boldsymbol{\theta}$ that are estimated from the data by minimizing a certain error function giving the goodness of fit with respect to $\boldsymbol{\theta}$. Often used for the error function is the root mean squared errors RMSE:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \varepsilon^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} [\boldsymbol{t}_i - f(\boldsymbol{x}_i, \boldsymbol{\theta})]^2} \tag{14}$$

with n the number of observations or data points.

For this experiment, $\boldsymbol{x}_i$ is the vector with predictors and $\boldsymbol{t}_i$ the vector with targets for the i^{th} tire assessment, for two drivers each having performed seven assessments, this results in a total of seven data points per driver. Having this data set available, two main problems arise when performing classical regression approaches: there is limited a priori knowledge available and this is small data set. These problems are explained below.

Limited a Priori Knowledge

If there is a priori knowledge available, this can be used for choice of the model structure. The first choice for modeling is to implement knowledge about a physical relationship using a mechanistic model where the parameters are estimated with regression. The advantage of such a physical model is that it can be applied for regression within the data range, but could also be applied for extrapolation, *i.e.*, outside the data range, as long as the physical relationships hold. For example, modeling a vehicle suspension by using a spring and damper model and using their physical relationship. Further a priori knowledge, like the model will only be used in the linear region of the spring and damper, can result in a model structure of a linear second order system where the parameters to be estimated can be the spring constant, damper coefficient and/or mass. For the relationship between the predictors and targets of this experiment, there is no physical relationship or other a priori knowledge for the model choice available, so this approach cannot be used.

Many regression methods are based on assumptions of the data or the relationship. If these assumptions are not met, the performance of the method is impaired. A well-known example is a linear regression model

$$t_i = \theta_1 f_1(x_{i1}) + \theta_2 f_2(x_{i2}) + \cdots + \theta_k f_k(x_{ik}) + \varepsilon_i \quad i = 1, \dots, n \qquad (15)$$

with n the number of data points and k the number of predictors. This model assumes that the target variable is a linear combination of the model parameters and (a function of) the predictor variables. Other assumptions could be no collinearity in the predictors, independence of errors and constant error variance. Although methods have been developed to bypass some of the assumptions, these methods often require many more data points and have themselves additional assumptions (Field, 2009). As for the relationship between predictors and targets for this research there is very little a priori knowledge available and therefore many assumptions cannot be made. An 'assumption free model' would therefore be preferable.

Limited Data Set

The second mentioned problem with the available data set is the few observations or data points available. For a regression model, there are seven data points available with 29 metrics that can be used as predictors for each driver. A linear regression model for this data set is given by Eq. (15), with $n = 7$. In this model, there are k parameters θ to be estimated. This system of equations defining the regression model is underdetermined when using more

than seven predictors, implying that the k parameters $\theta_j, j = 1, 2, \ldots, k$ cannot be estimated uniquely. Using extensions of this linear model as piecewise regression models like B-splines does not offer a solution; these methods also require a larger data set. A commonly used solution is to decrease the number of predictors, by leaving the less contributing predictors out or combining several predictors in one new predictor, for example using principal component analysis or factor analysis (Hyötyniemi, 2003; Johansson, 2006). These methods have the risk of obscuring the relationship between predictors and targets, by using composed variables rather than the more easily interpretable metrics. For this experimental data there is little a priori information available about how much a predictor contributes to the outcome and if this is the same for both drivers. In addition, to end up with an overdetermined system, which enables the use of most classical regression analysis, the number of predictors should be (far) less than seven. This requires eliminating or combining a large number of predictors with the risk of losing information and making the relationship less clear for interpretation.

6.2.3 Chosen Regression Method

To model a certain unknown function between predictors and targets a regression method that does not require strong assumptions or a large data set is a GRNN, which is a specific ANN. A detailed motivation for choosing this specific method requires more ANN background knowledge, therefore, this is described in the next section, where an introduction to ANN's is given first. In short, it does not require making assumptions about the data or the relationship; therefore, it does not require choosing a regression model. Instead, it empirically determines the appropriate relationship from the data as a probability density function. Additionally, it does not need a large data set to perform well. Although more data points will improve its performance, even with a limited data set a GRNN is able to converge to the underlying (linear or nonlinear) regression hypersurface, with a smooth transition between data points (Specht, 1991).

6.3 Generalized Regression Neural Network

In this section, a GRNN is explained by first introducing ANN in section 6.3.1. The motivation for using a specific ANN, a GRNN, is given in section 6.3.2. Section 6.3.3 gives an overview of the components of a GRNN.

6.3.1 Introduction to Artificial Neural Networks

An ANN is a biologically inspired computational model that consists of relatively simple processing elements, called 'neurons', the connections between them as well as the training or learning method (Kasabov, 1996). ANN's belong to problem-solving methods that are not derived from classical theories, but have analogy with biological reasoning and problem solving (Koivo, 2001). ANN's are based on the low, subsymbolic level of the brain and are based on the

physical operating of the brain. The human brain has an estimated 10^{11} neurons, which are interconnected with an estimated 10^{15} links (Munakata, 1998). Especially these connections and the strength of these connections play an important role and contain the 'knowledge' of the system. Although ANN's are far from even approaching the complexity of a real brain, several to dozens interconnected neurons in an ANN can already perform complex tasks, which are very difficult to perform by traditional computing. Two main advantages of ANN are the way they can model (hidden) relationships between input and output, without making assumptions about the data or the underlying relationships and the ability to learn from examples. This is especially useful when for a specific problem input and output data are present, but no or limited knowledge is available about the underlying relationship that maps the input to the output (a priori knowledge). An ANN is able to learn this underlying relationship from observations, combinations of input and output data, which in ANN terminology are referred to as 'examples'. These examples change the strength of the connections between the neurons in the ANN, capturing this relationship. Because of this feature, ANN's are also referred to as being 'non-algorithmic'. Furthermore, for several applications the 'graceful degradation' feature of ANN's is important, meaning that if there is a mistake in the ANN or input data are not complete or accurate, it does not shut down totally, but just performs less good. This is analogous to the working of the human brain. An illustration of this is reading a sentence where the vowels are left out, so you can only read the consonants. For example: "Cn y rd ths sntnc?" If not too complicated, you are able to read the sentence, but probably not as fast as you would normally do.

The development of ANN's started with the first model of a neuron in 1943 (McCulloch & Pitts, 1943) and has grown out to a nonlinear statistical data modeling tool for performing complex tasks, like classification, function approximation and control in various application fields (Prokhorov, 2008).

The artificial neuron

Fig. 52 illustrates the functioning of an artificial neuron with n inputs and one output. Every input $x_i, i = 1, \dots, n$ is multiplied with a certain weight $w_i,\ i = 1, \dots, n$ and then combined in a summing function. The result is fed into an activation function φ, giving the output, also called 'activation', y of the neuron:

$$y(\boldsymbol{x}) = \varphi\left(\sum_{j=1}^{n} w_j x_j\right) \tag{16}$$

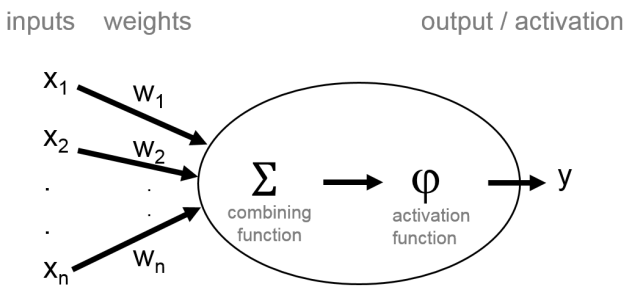


Figure 52. An artificial neuron.

Various functions can be used as activation functions; often they are nonlinear and bounded, like threshold functions, sigmoid functions or radial basis functions.

Network Architectures

Neurons can be combined, where more than one neuron can be connected to the inputs and where the output of a neuron can serve as input for other neurons. This results in different ANN architectures. In this introduction one specific ANN is briefly explained, a multilayer feedforward ANN with error backpropagation. Although this is not the ANN used for this research, it is by far the most widely used ANN and several issues discussed here, will later be referred to. Fig. 53 shows a simple example of a multilayer feedforward ANN. The neurons are arranged in layers and the information is processed in a feedforward way where every output is connected to neurons in the next layer, there are no feedback loops. Layers between the input and output layer are called hidden layers.

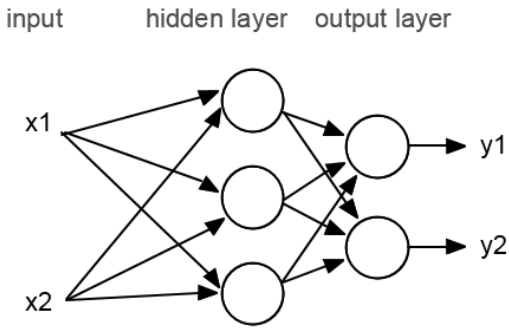


Figure 53. Multilayer feedforward ANN.

Not explicitly shown in the figure, but every input or connection has a weight value with which the input value is multiplied. Knowledge is captured in an ANN by adaptation of these connection weights. This is called 'learning' or ‘training’

of the ANN and various learning methods exists. An often used learning method is supervised learning, in which an ANN learns from examples, being a combination of input and output data. Based on these examples the ANN will adjust its connection weights to minimize the error between the actual output of the examples, the targets, and the calculated ANN output, the estimation of the tagets. The most widely used combination of ANN and training method is a feedforward, supervised learning ANN that uses the error back propagation algorithm for training (Kasabov, 1996). In this algorithm, the error of the output is propagated back to the input while adjusting the weights in order to reduce the error. This process is iterated many times by presenting the examples several times to the ANN, which can be up to thousands of times to converge to the desired solution. Depending on the data, the size, the architecture of the ANN and the hardware used, this training phase can take from seconds to days, but once the network is properly trained, it's connection weights are frozen. The network enters the next 'prediction phase', where it is ready for usage: making predictions of targets, based on new, unknown predictors. In this phase, it can operate fast and is suitable for real time applications.

For proper learning of this kind of ANN, many examples with enough variation must be available for training to let the ANN learn the underlying regression surface. Overfitting is a common problem in ANN training. Assuming that data have random noise, the network should ignore this noise and fit the underlying function. If overfitting takes place, the total network error becomes very small, the ANN learned the examples very well, but has not captured the underlying function. This is often overcome by dividing the available data set in a so-called training set and validation set. During training with the training set, the ANN is tested with the validation set (which is not used for training the ANN). When the total network error of the validation set stops decreasing or increases, the training phase is ended and it is assumed the ANN learned the underlying function without overfitting.

Although ANN's are often referred to as 'assumptions free methods', because assumptions about the data and regression model are not necessary, this does not release the user from using a prior knowledge. Application of an ANN also involves several choices to be made, often depending on the data and the expected behavior, like selection of network architecture with the number of hidden layers and the number of neurons used in each layer, feedforward or feedback, training and validation methods, data preparation. These are all important decisions that will influence the results that can be achieved. For example, if you want to model a discontinuous function with a feedforward ANN, at least two hidden layers are necessary (Russell, 2009).

6.3.2 Motivation for using a GRNN

There are two main problems for using classical regression methods with the available data set: there is little a priori knowledge available and the data set is limited in size, as explained in section 6.2. To use an ANN for regression could largely bypass the first problem, because an ANN does not require choosing a

regression model on beforehand, as described in the previous section. Nevertheless, the limited data set prevents the use of the widely used multilayer feedforward ANN with error backpropagation, as this would need at least a few hundred data points for proper learning (Kasabov, 1996). A GRNN is an ANN that is designed for regression and can also perform well with a small data set available for training (Specht, 1991). In addition, it is able to converge to the underlying (linear or nonlinear) regression hypersurface, with a smooth transition between data points (Specht, 1991).

Next to these two main reasons for using a GRNN, there are additional characteristics that make the use of the GRNN advantageous in general compared to the most used ANN, the feedforward ANN with the error backpropagation. Often, a much smaller data set is sufficient for a GRNN to accomplish the same performance. The architecture of the GRNN is relatively simple, having only one hidden layer in a feedforward network structure. This narrows down the many choices that must be made when designing the GRNN.

In addition, it has a one-pass learning phase, which means that no iterations are done, which makes the construction of a GRNN very fast. In addition, no division of the data in a training and a validation set has to be done, which for small data sets would reduce the training data set even more. Both the architecture and the one pass learning phase are explained in more detail in the next sections.

A GRNN is less suitable for very large data sets (*e.g.*, over several thousand data points) and for approximating discontinuous functions, but both conditions do not apply for this research.

6.3.3 GRNN Components, One-Pass Learning Phase and Prediction Phase

A GRNN belongs to the ANN category of Radial Basis Networks; these ANN's have neurons with a radial basis function as activation function.

Radial Basis Function

A radial basis function is a function whose value depends only on the distance from a center, defined by

$$r(\boldsymbol{x}, \boldsymbol{c}) = \|\boldsymbol{x} - \boldsymbol{c}\| \tag{17}$$

with r the Euclidian distance between $\boldsymbol{c}$ the (one or multidimensional) center point and $\boldsymbol{x}$ the input vector as multidimensional data point.

Several radial basis functions can be applied for the radial basis neurons as activation function, the radial basis function used here is

$$\varphi(r) = exp(-r^2) \tag{18}$$

Fig. 54 shows a graph of this activation function. If $r = 0$, the data point $\boldsymbol{x}$ is then positioned exactly on the center point $\boldsymbol{c}$, the output of the activation

function is at its maximum output of 1. The further the data point $\boldsymbol{x}$ is positioned from the center point $\boldsymbol{c}$, the larger r becomes and the activation function output decreases. The activation function output is half its maximum value of 0.5 when $r = \sqrt{ln2} = \pm 0.83$

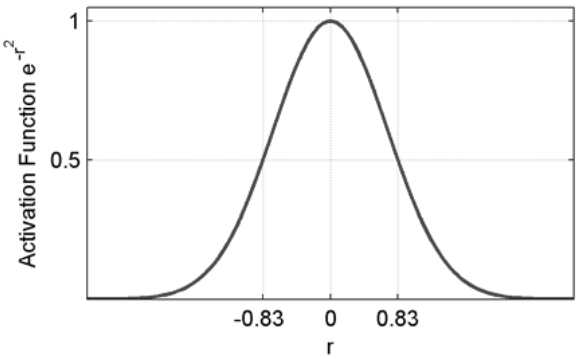


Figure 54. Radial basis function that is used as activation function in the hidden neurons.

GRNN Network Architecture

Fig. 55 shows the GRNN network architecture. It has a feedforward structure with one hidden layer. The architecture of the GRNN is fully determined by the examples used, as explained hereafter. After adding all the examples, the GRNN has 'learned' and is ready for application in the prediction phase, to predict the outcomes, based on new inputs. Because of this principle, a GRNN is said to have a one-pass learning phase.

Each hidden neuron represents one example or data point used for training (or constructing) the network, making the number of hidden neurons equal to the number of training data points. All the inputs are connected to all the neurons in the hidden layer. These connections have weights, called input weights. Each input weight is equal to the input value of the data point that the hidden neuron represents. All the neurons in the hidden layer are connected to all the neurons in the output layer. Each output neuron represents one output of a data point used for training the network, making the number of output neurons equal to the number of data point outputs. These connections also have various weights, called output weights. Each output weight is equal to the output value of the data point that the hidden neuron represents.

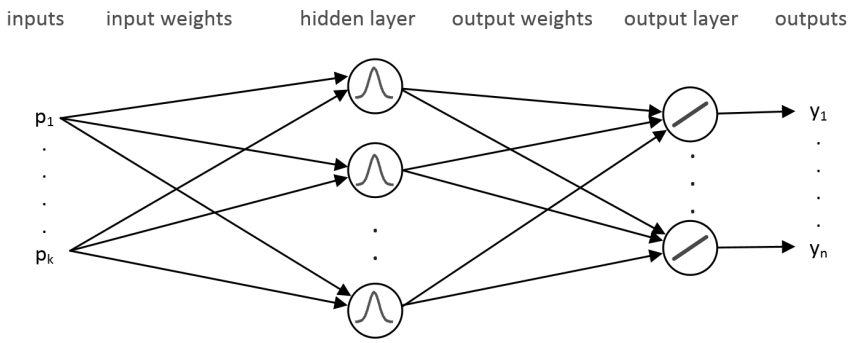


Figure 55. General architecture of a GRNN.

The hidden layer neurons have a radial basis function as activation function; the output layer neurons have a linear activation function. Both types of neurons are explained next.

Hidden Radial Basis Neuron

Fig. 56 shows a radial basis neuron in the hidden layer of the GRNN. This hidden neuron represents one of the examples for training/constructing the GRNN.

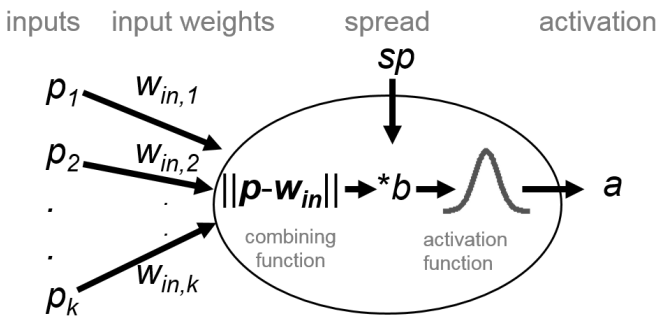


Figure 56. A radial basis neuron in the hidden layer.

This radial basis neuron has k inputs or predictors, the input vector $\boldsymbol{p}$. Each connection between one predictor and the neuron has an input weight $w_{in,j}$, $j = 1, \ldots, k$, giving k values in the input weight vector $\boldsymbol{w_{in}}$. This weight vector equals the input vector of the example represented by the neuron. The combining function for this neuron calculates the distance r between the two k-dimensional points represented by the weight vector $\boldsymbol{w_{in}}$ and the input vector $\boldsymbol{p}$

$$r(\boldsymbol{p}) = \|\boldsymbol{p} - \boldsymbol{w_{in}}\| = \sqrt{(p_1 - w_{\text{in},1})^2 + .. + (p_k - w_{in,k})^2} \qquad (19)$$

The distance r therefore represents the similarity with the example. The smaller the distance r, the more the given input resembles the example input, for the limit case, where $r = 0$ the input being equal to the example input. The calculated scalar output r is then multiplied with a bias value b before it is fed into the radial basis activation function φ of Eq. (18), giving the output or activation a of the neuron

$$a = \varphi(r \cdot b) \tag{20}$$

with bias value b

$$b = \frac{\sqrt{ln(2)}}{sp} \tag{21}$$

with sp the spread value, which is a parameter defined by the user. This spread value determines the distance at which the output of the radial basis activation function equals 0.5. It determines the sensitivity of the radial basis function. This can be visualized as making the radial basis function narrower or wider, as shown in Fig. 57.

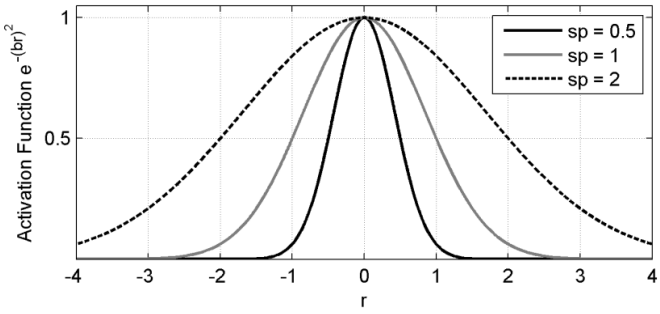


Figure 57. Three activation radial basis functions for different values of the spread sp.

In this figure, three radial basis functions are plotted, each with a different spread value sp. For the same distance r the activation radial basis function with the largest spread will produce the largest output value. A small value of spread results in an activation function that only produces an output significantly larger than 0 when $|r|$ is small, indicating that the input is quite similar to the example input. For $|r| = sp$ the activation function outputs 0.5.

Output Linear Neuron

Fig. 58 shows a linear neuron from the output layer of the GRNN.

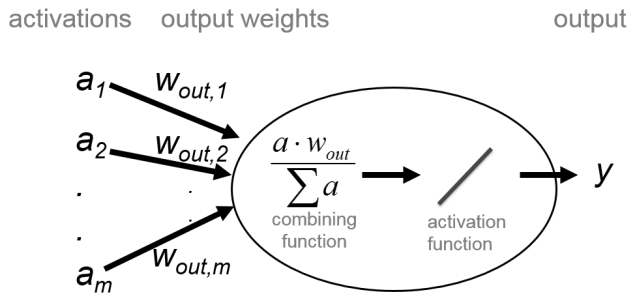


Figure 58. A linear neuron in the output layer.

This linear neuron has m inputs, which are the outputs or activations from the m neurons in the hidden layer, the activation vector $\boldsymbol{a}$. Each connection between an activation and the neuron has a weight $w_{out,j}$ $j = 1, \dots, m$, giving m values in the output weight vector $\boldsymbol{w_{out}}$. An output weight value is equal to the output value of the example, which the hidden neuron where the activation comes from, represents. The combining function for this neuron calculates the dot product of the weight vector $\boldsymbol{w_{out}}$ and the activation vector $\boldsymbol{a}$ and normalizes this result with the sum of activations. The activation function for this neuron is the identity function, so the output y is calculated by

$$y(\boldsymbol{a}) = \frac{\boldsymbol{a} \cdot \boldsymbol{w_{out}}}{\sum \boldsymbol{a}} = \frac{a_1 w_{\text{out},1} + \cdots + a_m w_{out,m}}{a_1 + \cdots + a_m} \qquad (22)$$

GRNN One-Pass Learning and Prediction Phase

The GRNN learning phase, which is its constructing phase, is straightforward and almost completely determined by the examples that are used for this constructing. The number of examples determines the number of hidden neurons; the input weights of each hidden neuron are equal to the input values of the example; the output weights are equal to the output values of the example. The only parameter is the spread value, which is relevant in the prediction phase of the GRNN.

In the prediction phase, when presented with an input vector, the GRNN will give an output vector, depending on this input vector. The more the input vector resembles the weight vector of a hidden neuron, representing the input vector of a 'learned' example, the higher the activation output of this hidden neuron. Subsequently, this hidden neuron high activation value is one input value for every output neuron, causing to weigh heavily the example output value, represented by the output weight, in the total output value of this output neuron.

For one input vector, this is done in parallel for all hidden neurons or examples, giving all their activation values to all output neurons. The actual output vector being a weighted average of all activation values multiplied with the output weights or examples output weight vectors. In general, the more an input vector resembles an example input, the more the output vector will resemble the example output. This does not imply that if a GRNN is presented

with the exact same input vector of one of the examples used for constructing the GRNN, the GRNN will output the exact same output vector of this example. This depends on the spread value and on how many similar examples are used.

The spread value determines how much 'similarity' of the examples is taken into account for constructing the output. The higher the spread value, the more examples are taken into account, the more the output will be a weighted average of more example outputs. For high spread values, all examples are included and the predicted output value will approach the mean value of the example outputs. For very small spread values, no similar examples are taken into account and the value of the output vector will be equal to the value of the example output. This learning and constructing phase is demonstrated for a simple example GRNN in the next section to demonstrate this theory.

Regardless the spread value, using a GRNN for prediction does not suffer from the mentioned limitations for prediction in section 6.2.1. As no regression model has to be chosen in advance, no inappropriate model can be chosen. Using a GRNN, the data itself deliver the ingredients for the model by positioning the radial basis functions on the data predictor points and scaling the activation functions with the target data points. The fitted curve will behave well between data points; its value stays between the data points target values and does not show extreme values as can be seen with fitting of higher order models. Also outside the data range, the curve does not show extreme values, it will approach the target value of the nearest neighbor, being the last data point in the range it exceeds.

Learning during Prediction Phase

During the prediction phase, when the GRNN is in use, it is easy to add new examples to the GRNN, when they become available. A new hidden neuron representing this example is added to the existing GRNN and this information is now included in future predictions. The more examples become available, the better the GRNN prediction will become.

Simple Example GRNN

A simple example is presented here for three data points of one predictor and one target value on which a regression curve is fitted with a GRNN. Assume we have three examples or data points for constructing the GRNN: (1, 1), (3, 3) and (4, 2). The corresponding predictor and target vectors are: $\boldsymbol{p} = [1, 3, 4]$ and $\boldsymbol{t} = [1, 3, 2]$. In the one-pass learning phase of a GRNN, the network is constructed by the data points. In the hidden layer, each neuron represents one data point by setting the input weight equal to the predictor value of that data point. This results in three hidden neurons. In the output layer, each neuron represents an output, which for this example results in one output neuron. The output weights from the hidden neurons to the output neuron are set to the data point target value. Fig. 59 shows the GRNN for this example with the input and output weights added.

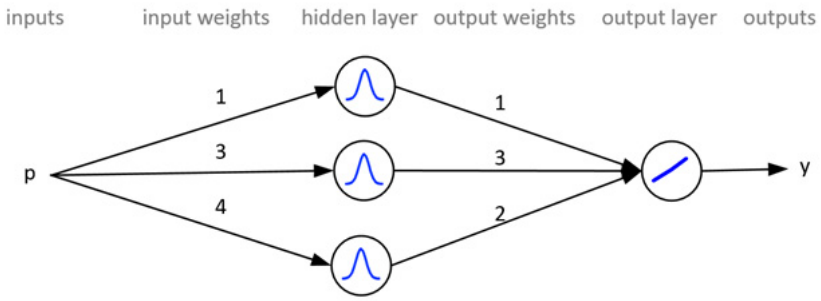


Figure 59. GRNN implementation for three data points (1, 1), (3, 3) and (4, 2).

The value for the parameter spread should be chosen to finalize the GRNN.

After this learning or constructing phase, the GRNN can be used for prediction in the prediction phase. For this example GRNN, if a new predictor value is presented as input, all three hidden neurons calculate the distance between this input value and its input weight. This distance is then fed into the radial basis function and, depending on the parameter spread, will produce an activation output of the hidden neuron. The more near this new input is to the data point represented by this hidden neuron, the closer the activation value is to 1. The activation outputs from the hidden neurons is then multiplied with the corresponding output weights and normalized with the sum of activation outputs in the output neuron. Fig. 60 shows the data points with the corresponding radial basis functions with a spread value of 1 and the resulting output y for this example GRNN.

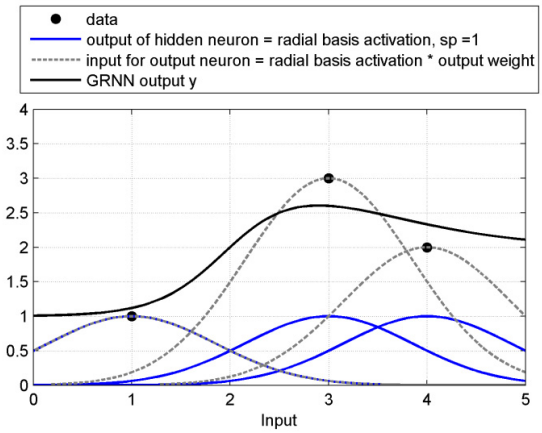


Figure 60. Data points, radial basis activation with a spread of 1, input for the output neuron and the GRNN output y.

The centers of the radial basis functions are positioned at the predictor values point of the data points, where the activation output is at its maximum of 1. Further from the center, the activation output decreases determined by the parameter spread. In this figure, the spread is equal to 1, meaning that the activation output will be 0.5 at a distance of 1 from the center. As can be seen, for this spread value, the radial basis functions have considerable overlap, which produces a smooth regression curve. To explain the calculations from input to output for prediction, we take as example input value $p = 4$. The hidden neuron representing data point (1,1) has a very small radial basis activation output, which we will take as 0 for this example. The hidden neuron representing data point (3, 3) has a distance of 1 with the input value, resulting in a radial basis activation value of 0.5. The hidden neuron representing data point (4,2) has the same input weight as the input value, the distance between the data point predictor value and this input is 0, resulting in an activation value of 1. In the output layer, the output value y is calculated by multiplying the activations with the corresponding output weights, being the target values of the data points, which is then divided by the sum of activations. For this example, this gives:

$$y(4) = \frac{0 * 1 + 0.5 * 3 + 1 * 2}{0 + 0.5 + 1} = \frac{3.5}{1.5} = 2.3 \qquad (23)$$

In this way, for all inputs between 0 and 5 the GRNN output y or regression curve can be calculated and is plotted in Fig. 60.

If the value of spread is made smaller, the radial basis functions will overlap less, resulting in a less smooth regression curve. In the extreme case, for a very small spread value, the radial basis functions have no significant overlap. This results in the so-called 'nearest neighbor value' for y, in which the GRNN output y will have the same value as the target value of the closest input data point. Fig. 61 shows the example GRNN, but now with a spread value of 0.1.

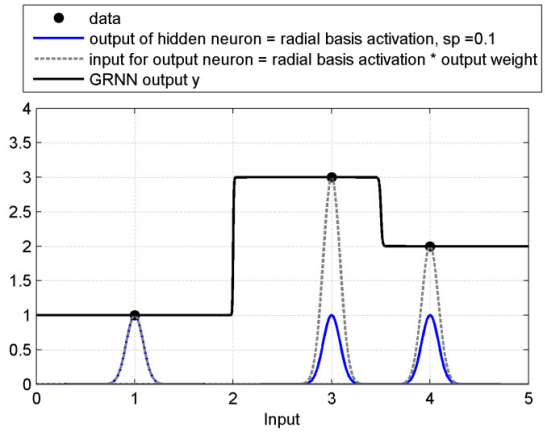


Figure 61. Data points, radial basis activation with a spread of 0.1, input for the output neuron and the GRNN output y.

If the spread is made larger, the radial basis functions overlap more; the output y becomes smoother as more data points influence its value. In the extreme case all the radial basis function overlap each other and the result is the mean value of the data points target values as output for all inputs. Fig. 62 shows the example GRNN with a spread value of 10.

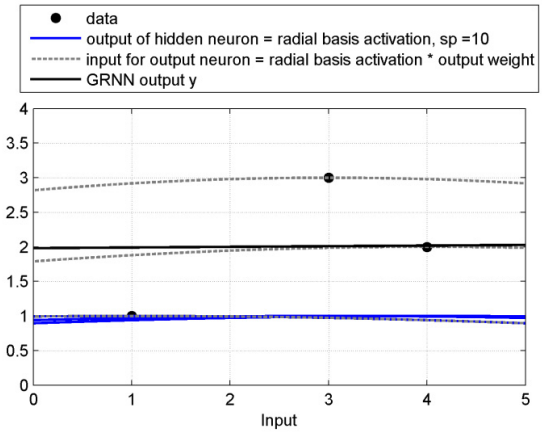


Figure 62. Data points, radial basis activation with a spread of 10, input for the output neuron and the GRNN output y.

As the only parameter of the GRNN, the spread value s determines where the regression curve lies between the nearest neighbor model and the mean value model. A priori knowledge about the data and the underlying behavior can help in determining this spread value. If the data contain much noise, implying the data points do not have to be fitted exactly, and/or the underlying behavior is assumed smooth, the spread value can be made larger. Fig. 63 shows three output curves for this example GRNN for different spread values. It shows that in this case presenting the input value of an example used to construct the GRNN, for $s = 0.5$ will give an output value almost equal to the example output value. For larger values of s, the output values of one or more nearby examples are taken into account, changing the output value from the example output value, making the output curve smoother.

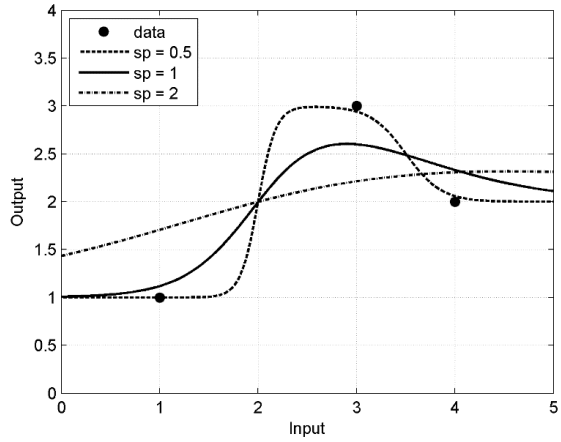


Figure 63. Data points and GRNN output y for several spread values.

6.4 GRNN for Handling Assessment

In this section, the GRNN is explained that was developed for modeling the relationship between the predictors, the (sub) set of metrics derived from the vehicle dynamics measurements, and the targets, the (sub) set of subjective evaluation scores on handling aspects, as described in the introduction of this chapter and in section 6.1.

6.4.1 Learning phase

In the learning phase, the GRNN for every driver is constructed as described in section 6.3.3. The examples used for constructing the GRNN are (a subset of) the assessments. Which assessments are chosen for construction is described in section 6.4.2. Fig. 64 gives an overview of the basic GRNN as result of the learning phase for one driver. The number of inputs or predictors is determined by the number of metrics that is used. Each input weight between a predictor and a hidden neuron is set equal to the (normalized) predictor value of the assessment represented by the hidden neuron. Each output weight between a hidden neuron and an output neuron is set equal to the (standardized) score of the assessment represented by the hidden neuron for that output aspect represented by the output neuron.

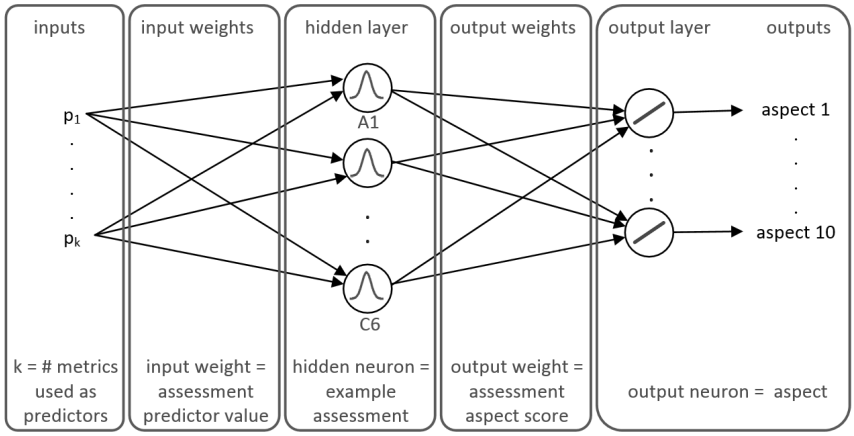


Figure 64. GRNN for handling assessment.

Previous research (Monsma & Defilippi, 2011) showed that using all metrics as predictors did not result in the best performance. This is because not all predictors have a similar contribution to the performance of the GRNN. Apparently, by adding all predictors, some of them decrease the predicting performance of the GRNN. Using less than all predictors for constructing the GRNN implies that the number of predictors must be defined and which metrics will be selected for this number of predictors.

The number of predictors to use for the GRNN will be a model parameter, which can be optimized for best performance. This will be explained later.

In the previous research, the selection of metrics to be used as predictors was done based on the strength of the linear relationship between a metric and the output 'aspect 10 overall judgement' given by Pearson's correlation coefficient r (Field, 2009), calculated for a group of N values (x, y)

$$r = \frac{cov(x,y)}{s_x s_y} \tag{24}$$

with s the (sample) standard deviation defined in Eq. (12) and with the covariance between x and y:

$$cov(x,y) = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \overline{y})}{N-1} \tag{25}$$

with $\bar{x}$ and $\bar{y}$ the mean values. For $r = \pm 1$ the relationship is perfectly linear, for $r = 0$ there is no linear relationship. In general, values of ± 0.1 represent a weak linear correlation, ± 0.3 a medium and ± 0.5 a strong linear correlation, but the interpretation of r depends strongly on the context of the research (Field, 2009).

Aspect 10 is the most important handling aspect, which should include the other aspects. Therefore, the assumption is made that a metric is more useful as predictor if the correlation coefficient is closer to 1. As an example, two scatter plots of standardized metric – aspect 10 values for all assessments for driver mean are shown in Fig. 65. The plot on the left shows metric v_y-V, which has the highest value for Pearson's correlation coefficient r. The plot on the right shows metric a_y-V, which has the lowest r with the aspect 10 scores.

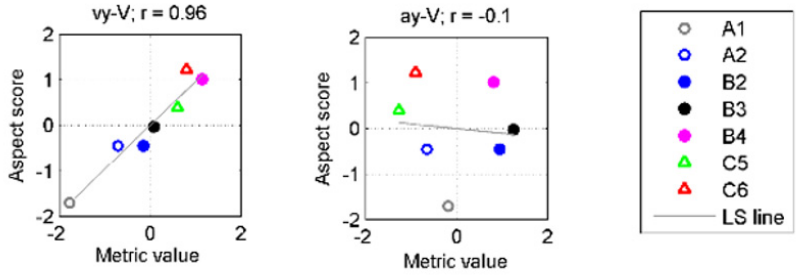


Figure 65. Two example scatter plots of standardized metric – aspect 10 values for all assessments for driver mean with the highest (left) and lowest (right) absolute r-value.

As can be seen from the plots, with a high absolute r, the metric value gives a good indication of the corresponding aspect score, as with a low absolute r, this is not the case. For the GRNN, all the metrics are sorted from highest to lowest absolute r-value and in this order, the metrics are added as predictors. As can be seen in Fig. 64, the selected predictors are used for predicting all 10 aspects, although only the correlation between the metrics with aspect 10 is used for the importance of the metrics as predictor.

For this research, the same method for selecting a metric as predictor based on the absolute r-value is used, but with the adjustment that for predicting a specific aspect, the predictors are selected based on the correlation with that specific aspect. This GRNN is visualized in Fig. 66 as 10 basic GRNN's in parallel, but each with only 1 output, being one of the 10 aspects. The number of hidden neurons is for each basic GRNN equal to the number of examples, the assessments, used.

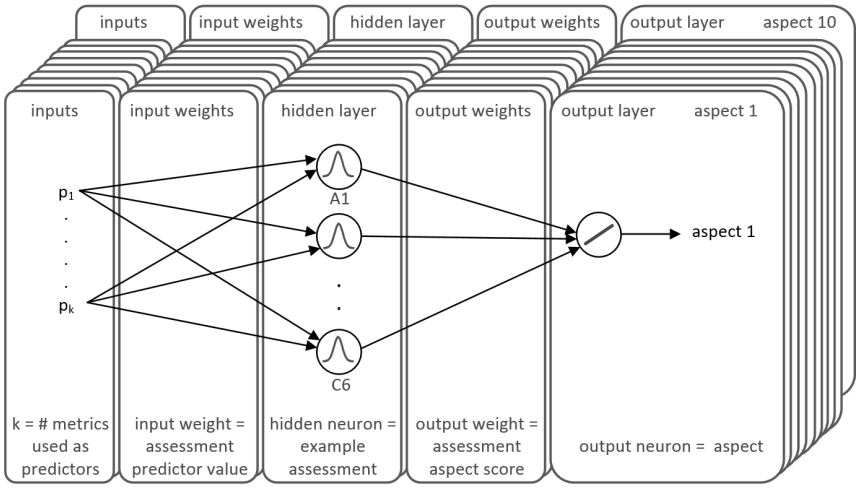


Figure 66. GRNN for handling assessment.

The number of predictors is, next to the spread value, a model parameter. Which metrics are used as predictors is not equal for the parallel GRNN's, as explained before; these are based on the correlation with the output, the aspect, of each basic GRNN.

Appendix 6 includes for every driver and for every aspect, the scatter plots of standardized metric – aspect values for all assessments sorted from highest to lowest absolute r-value.

6.4.2 Learning and Test Data Sets

The architecture of a GRNN is fully determined by the examples used for the one-pass-learning phase in which the GRNN is constructed. The examples which can be used for handling assessment, are the seven assessments per driver, with for every driver an 'own' GRNN, as explained in section 6.1. Constructing the GRNN based on all seven assessments would include as much information as possible, but this would leave no data, 'unknown' to the GRNN, for testing the prediction performance of the GRNN. Splitting the data in a separate learning set for constructing the GRNN and a separate test set also has its drawbacks. As explained in section 6.1.2, the subjective scores on the handling aspects for a certain tire are to a certain amount relative scores between batches, but taken for this research as absolute scores over all batches. Not using one or more assessments as examples in the learning phase of the GRNN could diminish a batch to one or no assessments, which reduces the performance of predicting similar examples. The way of splitting the data set would therefore have an unwanted significant influence on the model performance.

Leave-One-Out Cross Validation

A practical way of testing a model based on such a limited data set, is cross validation, with as special case the 'leave-one-out' method (Field, 2009;

Hyötyniemi, 2003). This method implies that one observation is left out of the data set to construct the model. After the model is constructed based on the remaining observations, the model is used to predict the left out observation outputs, based on the left out observation inputs. This procedure is repeated until all observations are once left out for validation. The performance of each model is evaluated by applying an error function between the predicted outputs and the actual outputs of the left out observation. The overall model performance is evaluated by taking the average error of all models.

This leave-one-out cross validation method is applied for this research but with the adjustment that the assessments with the most number of highest or lowest predictor values, referred to here as the 'extreme' assessments, are not left out for validation. As explained in section 6.2.1, prediction of targets should only be done with predictor values that lie inside the data range used for the model fitting. Although a GRNN will not behave badly outside the data range of the learned examples - it will approach the nearest neighbor model as explained in section 6.3.3 – the expected relative large error function value for prediction outside the data range could dominate the average performance calculation. This assumption is validated in section 6.5.

To determine which assessments are the extreme assessments regarding the predictor values, these predictor values are compared. As the metrics are all potential predictors, this is done for all the metrics by plotting their values and analyzing which assessments give in general the highest or lowest values. As an example, this is shown for two metrics for all assessments and for all drivers in Fig. 67. The plots show that for these metrics, assessments A1 give the lowest metric values and assessments C6 give in most cases the highest metric value.

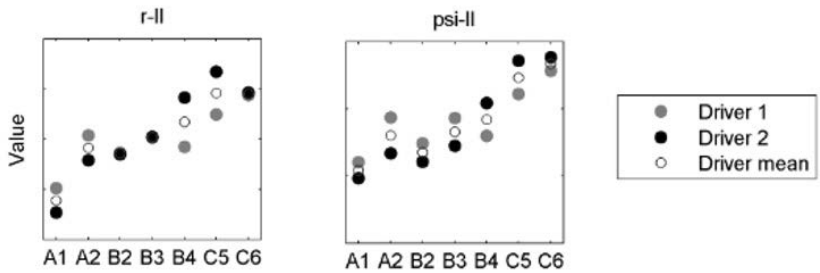


Figure 67. Example of two scatter plots of metric values for all assessments for all drivers.

In Appendix 4 these plots are given for all metrics and from these plots it can be seen that in most cases assessments A1 and C6 give the extreme metric values, which is also in accordance with the extreme output values, the aspect scores or targets. Consequently, these extreme assessments A1 and C6 are not left out for validation, but are always part of the training data set to prevent prediction out of the model fitting data range.

6.4.3 GRNN Prediction Performance

The prediction performance of the GRNN is tested by giving the GRNN inputs of the test assessment that is left out in the learning phase of the GRNN, so is unknown to the GRNN. Based on these inputs or predictors, being the metric values of the number of predictors used, the GRNN will predict the 10 aspect scores of the driver.

The GRNN for handling assessment has two parameters, the number of predictors used and the spread value. The prediction performance strongly depends on these parameter values. The leave-one-out cross validation should deliver the overall model performance of the GRNN prediction, which will be a function of the parameter values. The objective is to define the model parameter values or value ranges which result in the highest performance.

RMSE

As performance measure of the GRNN, the RMSE (Eq. (14)) between the actual aspect scores given by the driver and the GRNN predicted aspect scores is taken. The lower the RMSE value, the higher the prediction performance. The advantage of this performance measure is that large errors are penalized more than small errors due to the squaring. As this prediction should give an indication of the subjective assessment scores, small errors are negligible, as the drivers use a resolution of 0.25 points in their score, but large errors should be avoided. Compensating a large error for one aspect by many small errors for other aspects is less likely when using this measure. By taking the square root, the performance measure is in the same units as the aspect score and can be seen as an absolute performance measure.

Mean Model Performance

The most basic model is the mean model. As the name implies, this model predicts an output value by the mean of the other output values. This model is simple, easy to derive and often used as base line model for comparing other model performance, like the R square criterion, explained in the next section. The basic idea is that a model is only a useful model, if it performs better than the mean model. Therefore, the results of the GRNN prediction will also be analyzed in comparison with the mean model results.

For this handling assessment data, the mean model predicts the left-out target or aspect of the test assessment, by taking the mean of all the remaining example assessments target values, hence without the test target value. For example, for driver 1, aspect 1, the target values for the assessments A1 to C6 are respectively: 5, 6, 6, 6.75, 7.25, 7, 8 (see Appendix 5 for the assessment scores). Assume assessment A1 is left out as test assessment, the mean model will predict aspect 1 by the mean of the other 6 assessment target 1 values, resulting in the value 6.83. This prediction has an error of -1.83. Appendix 7 gives all the mean model predictions for every driver, for all aspects for all test assessments. As example, Fig. 68 shows the actual and mean model prediction aspect scores for test assessments A1, B3 and C6 for driver 1.

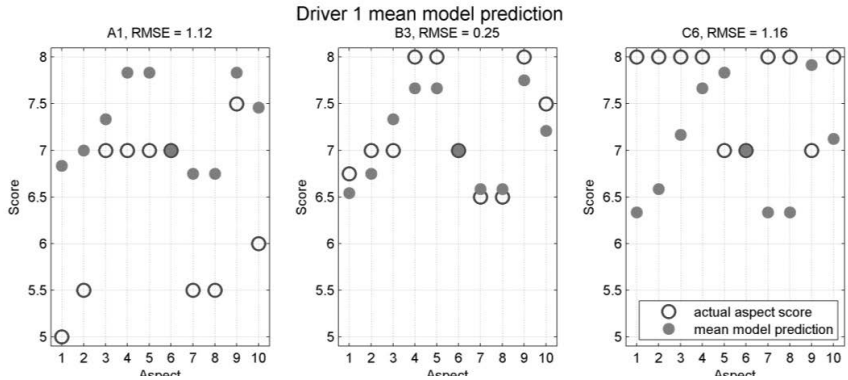


Figure 68. Plots of the actual and mean model prediction aspect scores for test assessments A1, B3 and C6 for driver 1.

The left hand plot shows that the mean model prediction for test assessment A1 is for all aspects, except aspect 6, too high, because A1 has the lowest actual aspect scores and these are not taken into account for calculating the mean values, as this is the left-out test assessment. The opposite is true for test assessment C6, shown in the right hand plot. This is as expected for the extreme assessments. For an intermediate assessment as shown in the middle plot for test assessment B3, the mean model is more accurate in predicting the aspect values. Aspect 6 was predicted without error for all test assessments, because for all assessments, this aspect scored a 7. The RMSE value shown in the title of each plot is calculated over all aspects for that test assessment, showing relatively high values for the extreme test assessments, as expected.

R squared criterion

The R squared criterion, often used for evaluating model data fit relative to the basic mean model (Hyötyniemi, 2003), was also considered as prediction performance measure for the GRNN, defined as

$$R^2 = 1 - \frac{SSE}{SST} \tag{26}$$

with the sum of squared errors

$$SSE = \sum_{n=1}^{10} (t_n - y_n)^2 \tag{27}$$

and the sum squared total

$$SST = \sum_{n=1}^{10} (t_n - \bar{t_n})^2 \tag{28}$$

with 10 the number of targets or aspects, t_n as nth target value, y_n as nth predicted target value and $\bar{t_n}$ as the mean model prediction as explained before.
This mean calculation is different than the commonly used definition where the mean value would be calculated over all 10 aspect scores of the test assessment, which is not relevant in this case. If the GRNN would predict the targets exactly, SSE would be equal to zero, giving an R^2 of 1. If the GRNN would predict the targets as the mean model, SSE would be equal to SST giving an R^2 of 0. R^2 values less than 0 indicate less good prediction than the mean model. As the mean is calculated over the target value of all assessments, except the test assessment, the mean model SSE value strongly depends on the test assessment, just as is shown for the RMSE. This makes this relative R^2 less suitable as GRNN performance measure, because it uses the strongly varying SSE as base line performance to compare the model performance. Therefore, the absolute RMSE value is used for the GRNN performance measure, which is compared to the absolute RMSE value of the mean model.

6.4.4 GRNN Learning during Prediction Phase

In this section, some remarks are given for the learning during the prediction phase of the GRNN for handling assessment, when new assessments become available or when metrics are added or removed.

As mentioned before, new examples can easily be added to a GRNN, including this information in future predictions, but also using the existing examples information. For this GRNN, the examples are new tire assessments, which can be extra assessments for the same tire or new, unknown to the GRNN, tires. Therefore, the GRNN receives an additional learning phase, in which the new examples are added. With more examples, the GRNN will learn more about the underlying relationship between predictors and handling aspects, the better the GRNN prediction will become.

Likewise, metrics can be added. If new measurements become available from which metrics can be derived, these can be added to the GRNN as inputs. If this metric will be used as predictor, the hidden neurons of existing examples have no input weights for this predictor, because they have no values for this new predictor. By setting these input weights to not-a-number value, this predictor will be ignored by this hidden neuron. The activation output of this hidden neuron will then be calculated as before, based on the other known predictors.

The same method can be used when metrics are removed, but for the existing examples still should be available. This can be because they are not relevant or are no longer derived as predictors for new assessments. This also accounts for aspects. New assessments with more or less predictors and/or more or less aspects than the existing examples can be incorporated in the GRNN, adding their information, while the existing assessments information is also kept.

For this research, many vehicle dynamics signals are measured and metrics are derived. The performance analysis of the leave-one-out cross validation will suggest an optimal number of predictors to be used for the prediction phase of the GRNN. Based on the actual predictors used during this prediction phase, signals, which are not necessary for the predictors, do not have to be measured

anymore during testing for GRNN use. In addition, analyzing the GRNN performance without predictors based on certain signals can reveal that those signal measurements can be removed with acceptable loss of prediction performance. This can be relevant when the measurements are difficult or expensive or when the metrics are not so relevant, *i.e.,* not or almost not used as predictor. This will be a cost benefits analysis.

6.5 Results and Discussion

The overall prediction performance of the GRNN for handling assessment is evaluated with leave-one-out cross validation, as explained in section 6.4.2. These results are described in section 6.5.2. Before this, the GRNN performance on the separate test assessments is analyzed in section 6.5.1, to see how the performance depends on the GRNN parameter values and on the test assessment used.

6.5.1 GRNN Prediction Performance on Separate Test Assessments

To validate the assumption that the GRNN prediction performance for the extreme assessments would be worse compared to the other test inputs, the GRNN prediction performances for all test assessments are compared. The GRNN is constructed based on the example assessments, but with the test assessment left out. The GRNN RMSE, calculated with the actual 10 aspect scores and the GRNN predicted 10 aspect scores, is shown in a surface plot as function of the GRNN parameter values, the number of predictors used and the spread value. These plots are presented and discussed in this section and for reference also included in Appendix 8.

Driver 1

Fig. 69 shows the results for driver 1. To compare the GRNN performance with the mean model performance, the RMSE value of the mean model prediction is also shown in the surface plot as grey solid horizontal lines representing the contours of the surface of the mean model RMSE value. As the mean model RMSE is a single value for a test assessment, this is shown as a flat surface in the GRNN surface plot. The mean model RMSE values per driver per test assessment are given in Appendix 7 in the last rows of Tables 29 to 31. The interpretation of this plot is that the GRNN performs better than the mean model, as long as the GRNN RMSE surface is below the flat mean model RMSE surface.

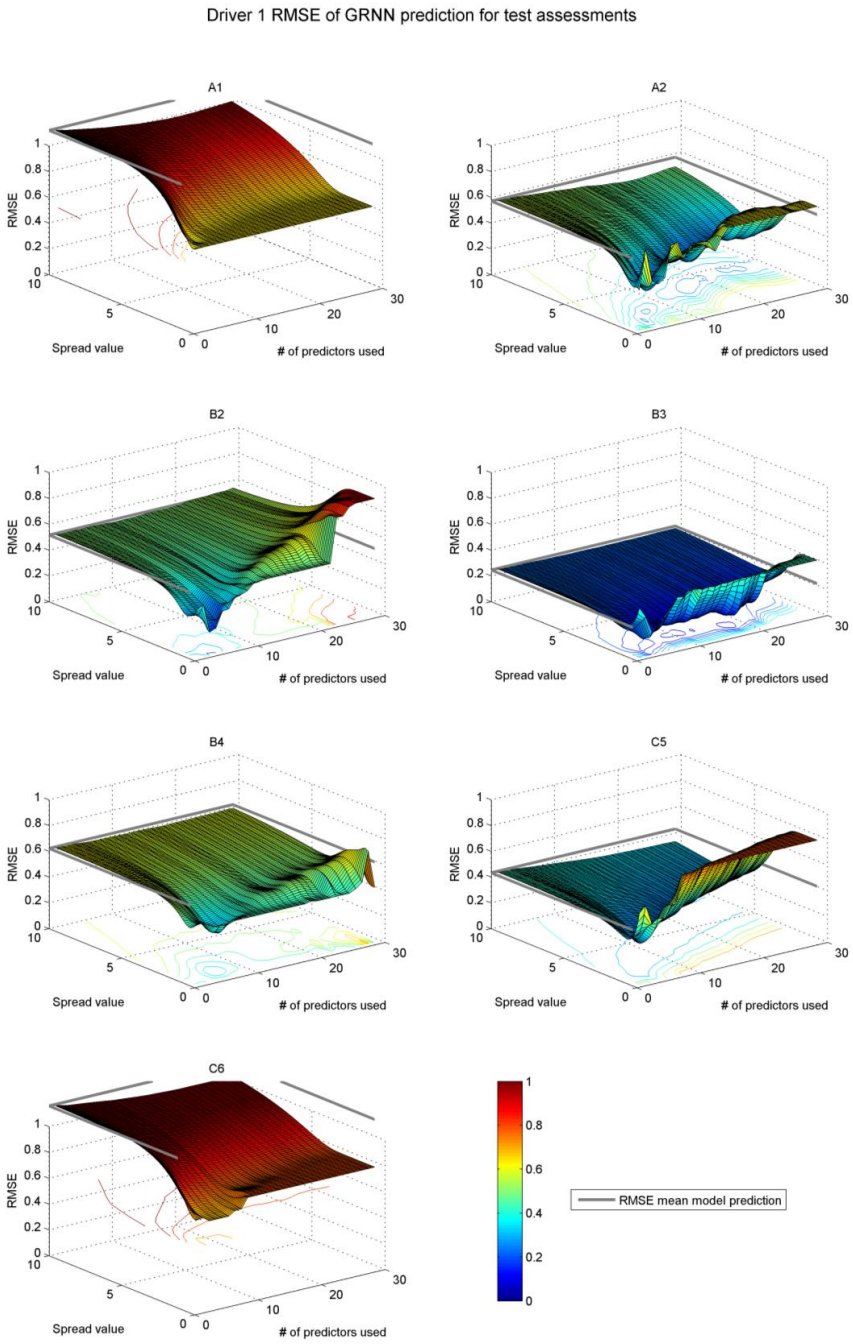


Figure 69. Driver 1 surface plots of the GRNN RMSE for all test input assessment as function of GRNN parameters spread value and number of predictors used.

Comparing the surface plots for the extreme test assessments A1 and C6 with the other test assessments it is clear that the RMSE values are higher. This

validates the assumption that the GRNN prediction performance for the extreme test assessments is worse compared to the other test assessments, due to the out of range data prediction.

As expected, for high spread values, the GRNN prediction approaches the mean model prediction, as explained in section 6.3.3. This is confirmed by the surface plots, where the surface for the high spread values approaches the mean model prediction RMSE, represented by the grey lines. This is confirmed by Fig. 70, where, as example, the GRNN predicted aspect scores are given for driver 1 test assessments A1, B3 and C6. Because the GRNN has a high spread of 10, the predicted aspect scores are equal or near the mean model scores.

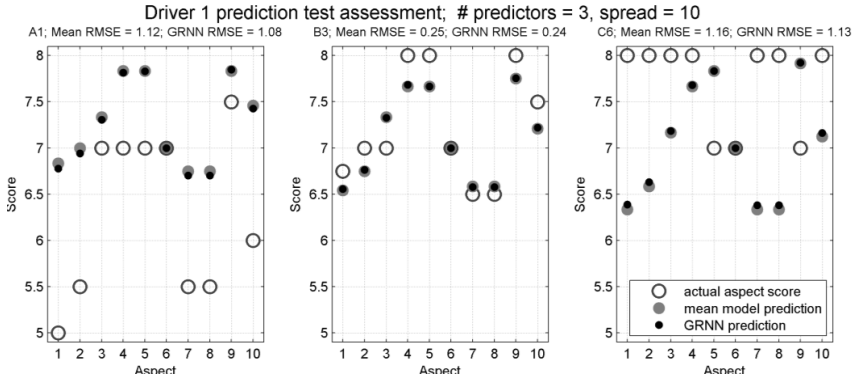


Figure 70. Plots of the actual, mean model and GRNN prediction aspect scores for test assessments A1, B3 and C6 for driver 1 with the GRNN having a high spread value of 10.

For low spread values, the GRNN prediction approaches the nearest neighbor model. The nearest neighbor of a test assessment varies, because it depends on the predictors used and with more than 1 predictor used, there can be more than one nearest neighbor. The outputs values, the predicted aspect scores, will be a combination of these nearest neighbor aspect scores. Therefore, this GRNN behavior for small spread values cannot easily be seen in the surface plots. The exception is test assessment A1, for almost all driver 1 predictors of A1, the predictor value of assessment A2 is the nearest neighbor. For small spread values, the GRNN will therefore predict for test assessment A1 as aspect scores the aspect scores of A2. This is shown in Fig. 71 for driver 1 test assessment 1 with the GRNN spread value of 0.1 the GRNN prediction of A1 are equal to the A2 actual aspect scores. The RMSE value of the actual A1 aspect scores and the predicted A2 aspect scores is 0.63, which is the value the surface plot of driver 1 of test assessment A1 is approaching for small spread values. Owing to being an extreme assessment, the surface plot shows a smooth transition from nearest neighbor model to mean model with increasing spread values.

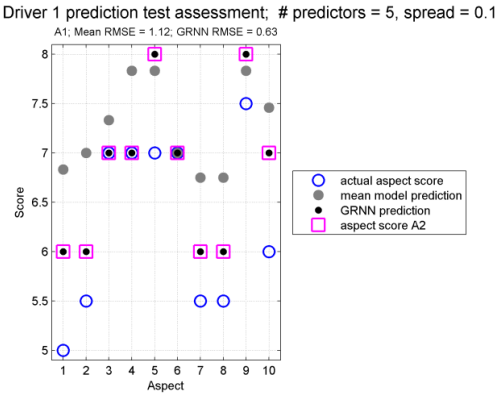


Figure 71. Plots of the nearest neighbor behavior of the GRNN, with a small spread value the GRNN predicts the nearest neighbor A2 aspect scores for A1.

For the intermediate test assessments, the performance of the GRNN is also approaching the nearest neighbor and mean model for the extreme spread values, but having several nearest neighbors that change with using different numbers of predictors, the performance is varying much more, shown by a less smooth surface.

Low spread values combined with a large number of predictors is a bad combination; the RMSE surfaces have high peaks here. The GRNN performance can even get worse than the mean model.

In general, the lowest RMSE values are found in a 'valley' running from low spread values combined with a small number of predictors to medium spread values combined with a large number of predictors (clearly seen from the contour plots in Appendix 8). For a driver's test assessment, the combination of number of predictors used and spread value that gives the lowest RMSE, is defined as the optimal GRNN prediction. These optimal GRNN predictions for driver 1 are shown in Fig. 72.

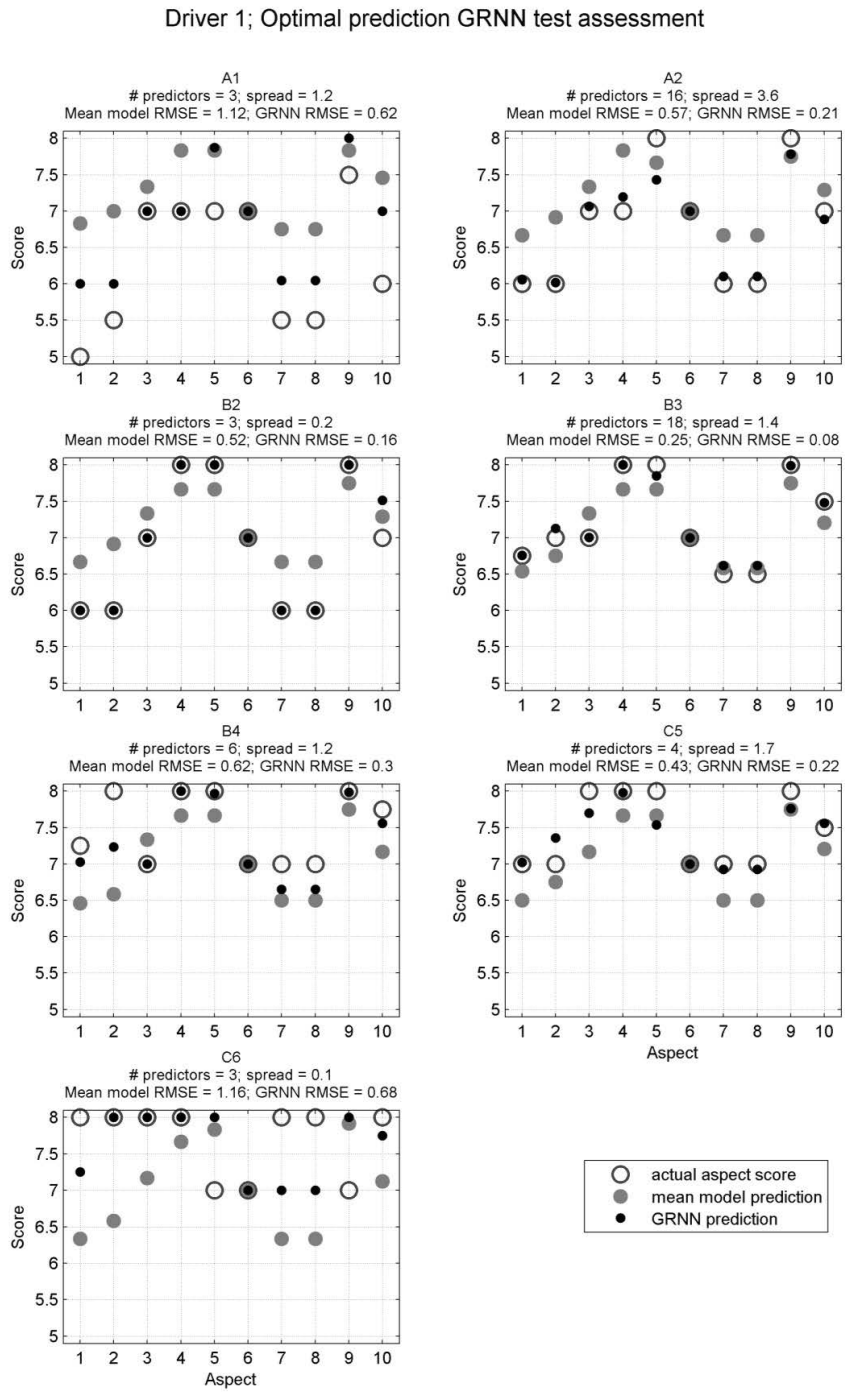


Figure 72. Driver 1 plots showing the actual, mean model and GRNN optimal prediction aspect score for all test input assessment.

This figure shows that the GRNN prediction is more accurate than the mean model prediction. Even for the extreme assessments, with prediction outside the data range, the GRNN accuracy is sustained in all the cases studies and it outperforms the mean model. As the drivers use a score resolution of 0.25, having a RMSE lower than this value for almost all intermediate tires, can be regarded as very good.

A comment about this performance must be made. Such an optimum can only be found because the test target values are known, and so the lowest RMSE value can be calculated. For predictions, where the target values are not known, the optimal values for number of predictors and spread must be estimated, which is the objective of the leave-one-out cross validation, explained in section 6.5.2.

Driver 2

Fig. 73 shows the results for driver 2.

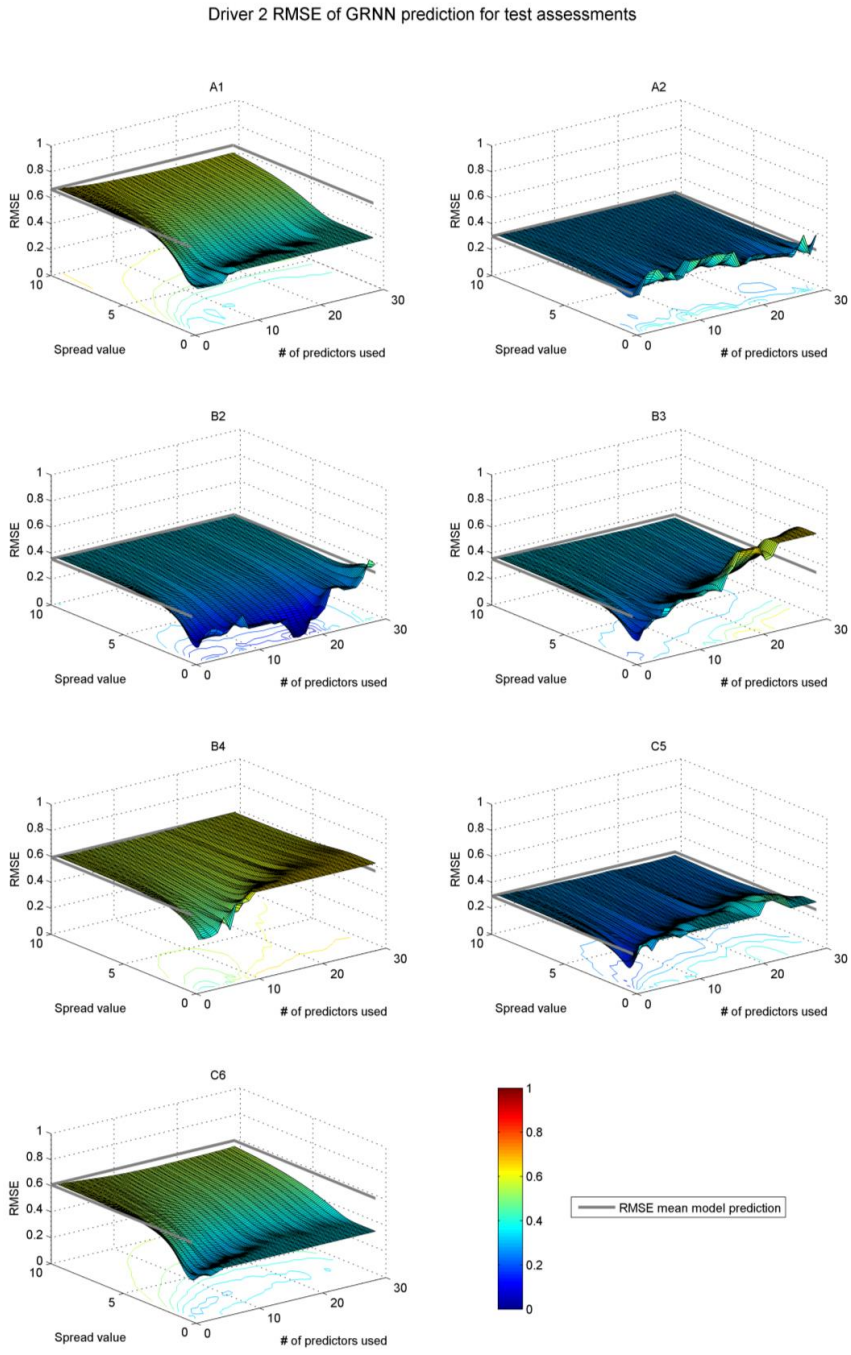


Figure 73. Driver 2 surface plots of the GRNN RMSE for all test input assessment as function of GRNN parameters spread value and number of predictors used.

For driver 2 the same high RMSE values can be seen for test assessments A1 and C6, but also for B4. This can be explained by the fact that the assessments were

rated absolute within one batch of one driver, but relative between batches and drivers, causing batch C to be rated lower as would be done if the ratings were absolute, as explained in section 6.1.2. This is more significant for driver 2 than for driver 1 as shown in Table 10 of that section. For the mean absolute score, assessment B4 was rated as best tire by driver 2. Seen from the 'predictors view', needed to prevent out of range prediction, A1 and C6 are the extreme assessments, but from the 'target view' for driver 2, assessment B4 is the extreme assessment. If this assessment is left out as learning example, to serve as test assessment, the GRNN learning examples miss the extreme target values of B4, so will predict too low B4 aspect scores, similar to the extreme assessments A1 and C6 as test input. The RMSE values for these test assessments are lower for driver 2 than for driver 1 for both the GRNN and mean model prediction. This is influenced by the smaller actual scoring range used by driver 2, *i.e.,* 2.25, compared to driver 1, *i.e.,* 3.00. Driver 1 scored aspects more extreme than driver 2, which for an absolute error function as RMSE will results in larger values.

The optimal GRNN predictions for driver 2 are shown in Fig. 74.

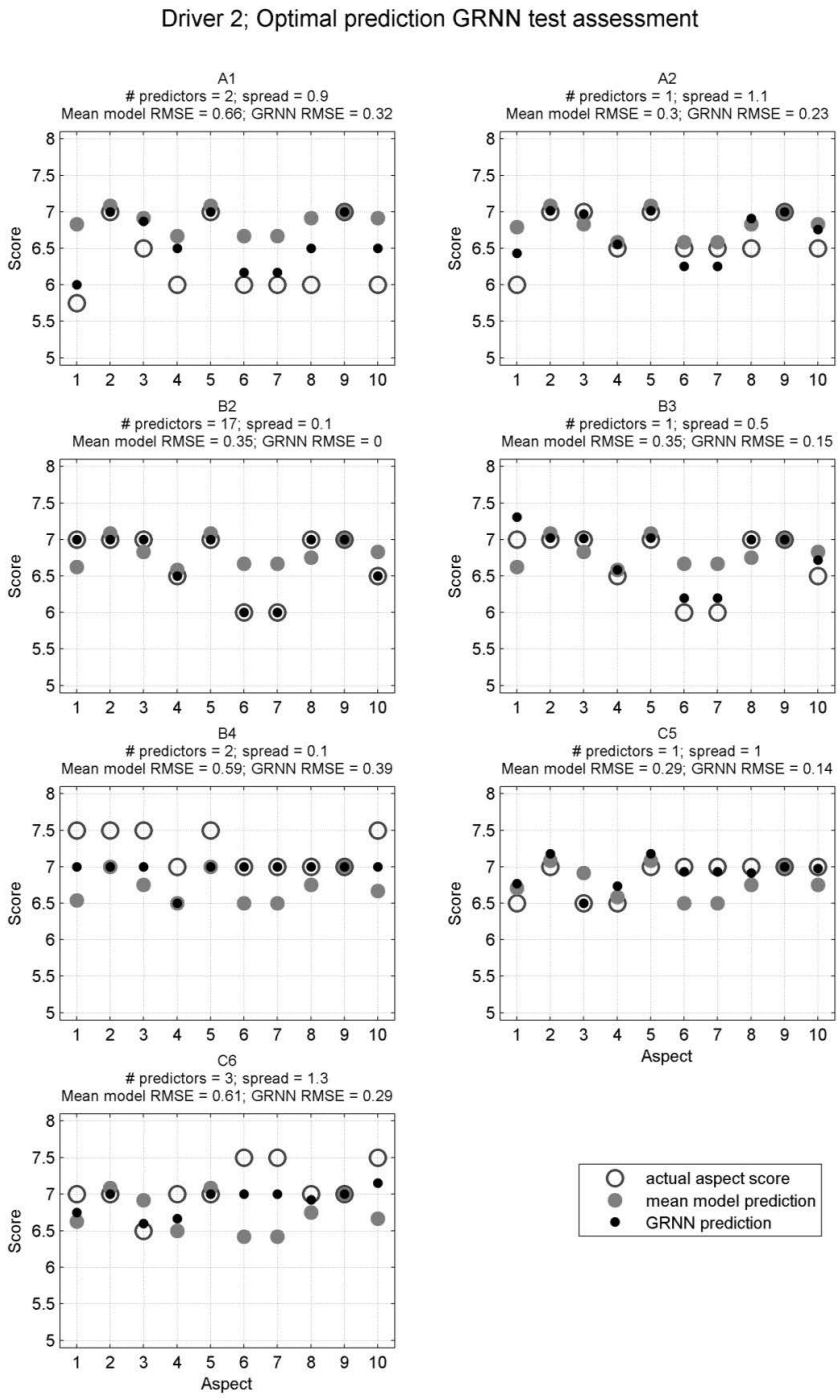


Figure 74. Driver 2 plots showing the actual, mean model and GRNN optimal prediction aspect score for all test input assessment.

These plots also show good GRNN prediction performance with all RMSE values around the drivers score resolution of 0.25, except for B4. Test assessment B4 has the highest RMSE of 0.39, which is in line with being the 'extreme target assessment' as explained before.

Driver mean
Fig. 75 shows the results for driver mean.

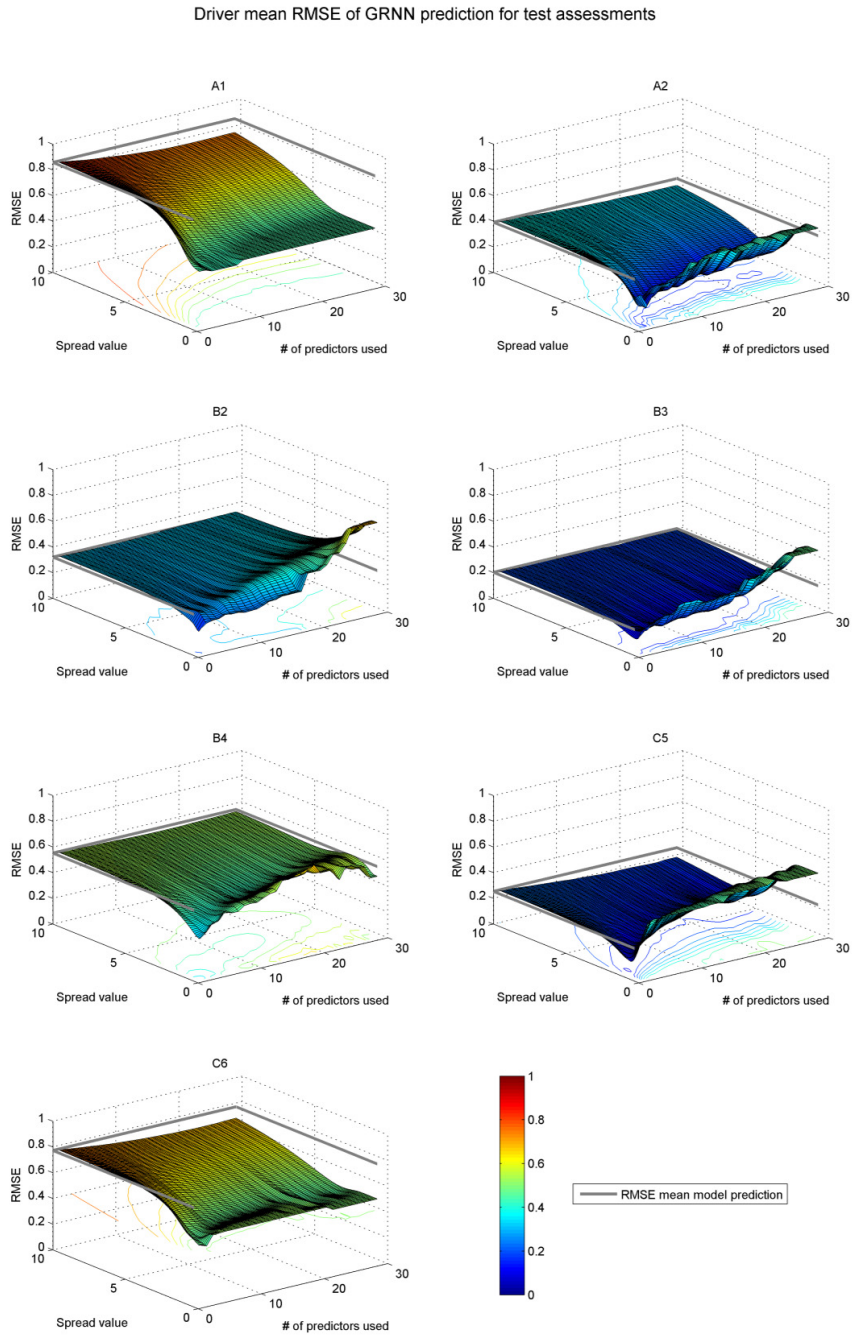


Figure 75. Driver mean surface plots of the GRNN RMSE for all test input assessment as function of GRNN parameters spread value and number of predictors used.

The RMSE surface plots for driver mean are smoother and less extreme, as expected, because of the smoothing effect of calculating a mean value. However, the general surface characteristics are still present, like the location and positioning of the 'optimal valley' and the nearest neighbor and mean model approaches for extreme spread values.

The optimal GRNN predictions for driver mean are shown in Fig. 76.

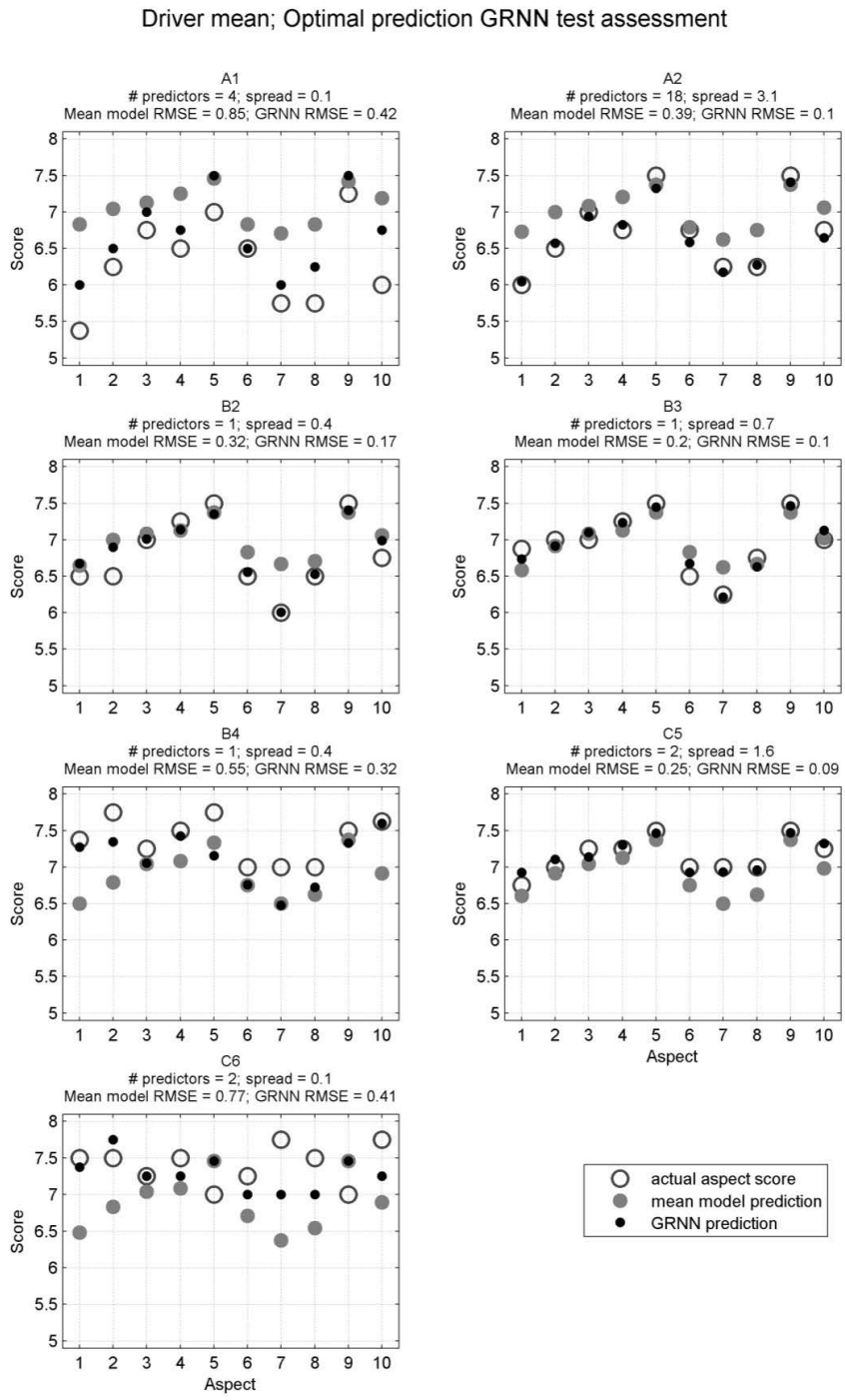


Figure 76. Driver mean plots showing the actual, mean model and GRNN optimal prediction aspect score for all test input assessment.

These plots also show good GRNN prediction performance with all RMSE values far lower than the drivers score resolution of 0.25, except for B4. Comparing all drivers, the GRNN performance for driver mean is the best. With the scoring range lying in between driver 1 and 2, this cannot be the reason of this. This best performance can be explained by the fact that calculating the mean values of the drivers smooths out extreme values, which could be undetected outliers that can have a negative influence on the performance.

Prediction performance on aspects

As stated in section 6.1, handling aspects 2 and 3, respectively stability while cornering with no throttle change and with throttle change, were determined while driving the circles. This behavior is not included in the metrics derived from the double lane changes, but will be predicted by the GRNN. Therefore, the assumption is that these aspects will not be predicted as well as the aspects related to driving the double lane change. Fig. 77 shows the mean of the absolute error between the actual aspect scores and the GRNN optimal predicted aspect scores.

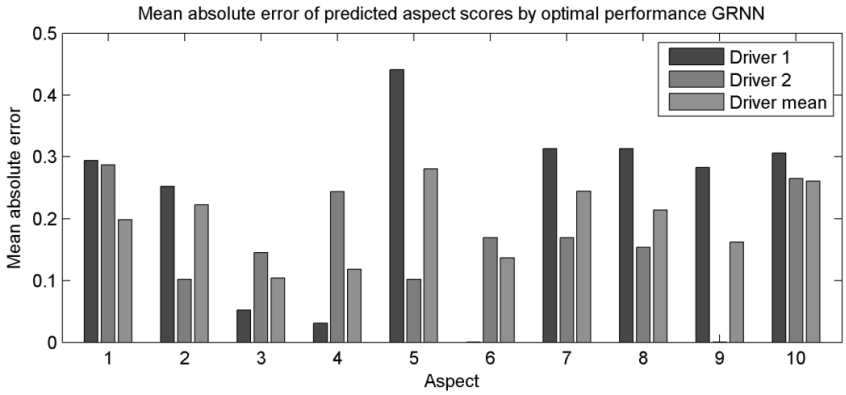


Figure 77. Mean absolute error between the actual aspect scores and the GRNN optimal predicted aspect scores for all drivers.

This figure shows that the assumption is invalidated; aspects 2 and 3 are not predicted less good than the other aspects. On the contrary, aspect 3 belongs to the best-predicted aspects. The GRNN is able to predict handling behavior for the total driven handling circuit, based only on the double lane change measurement.

6.5.2 GRNN Prediction Performance on Leave-One-Out Cross Validation

The GRNN for handling assessment is meant to be used for prediction of targets based on predictor values, where the target values are not known. For example, when a double lane change is driven during actual or virtual (simulated) testing and the data for deriving the predictors is available, but no subjective assessment is performed.

For prediction, the GRNN parameters, number of predictors used and spread, must be assigned a value. These values will be estimated by the values that result from the leave-one-out cross validation, that result in the lowest average error of the left out test assessments. Therefore, the surface plots and contour plots of the average RMSE of test assessments A2 to C5 are presented for all drivers in Fig. 78 and 79 .

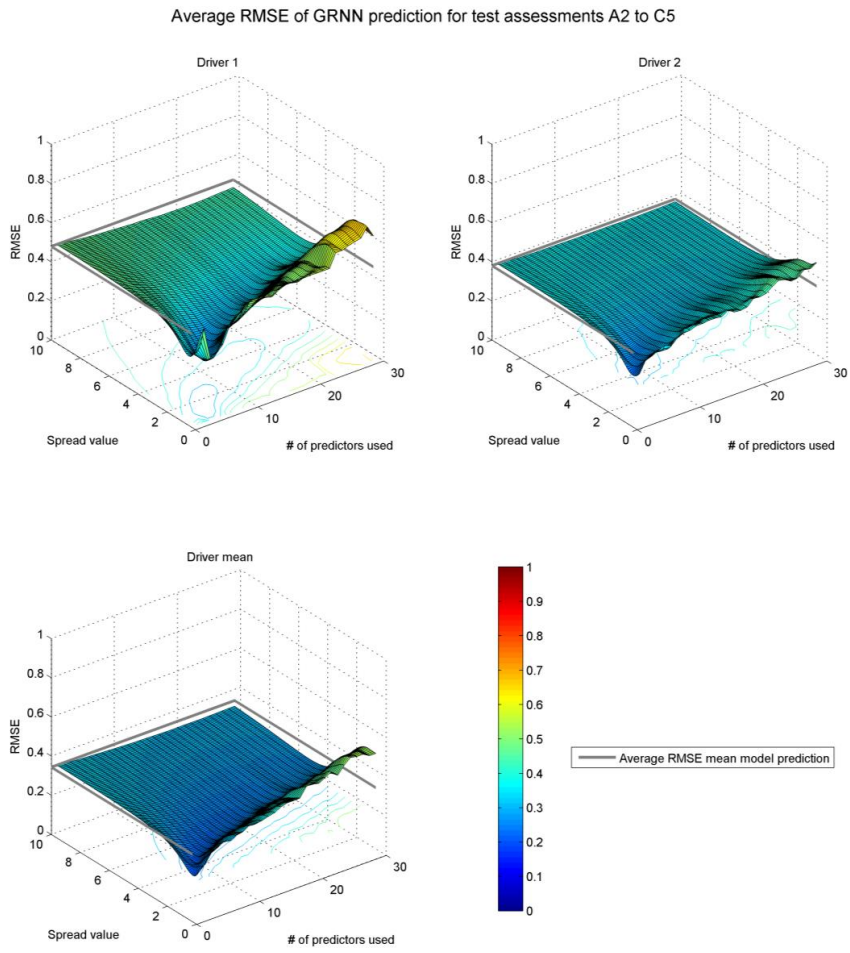


Figure 78. Result of the leave-one-out cross validation: surface plot of the average RMSE of GRNN prediction for test assessments A2 to C5 for all drivers.

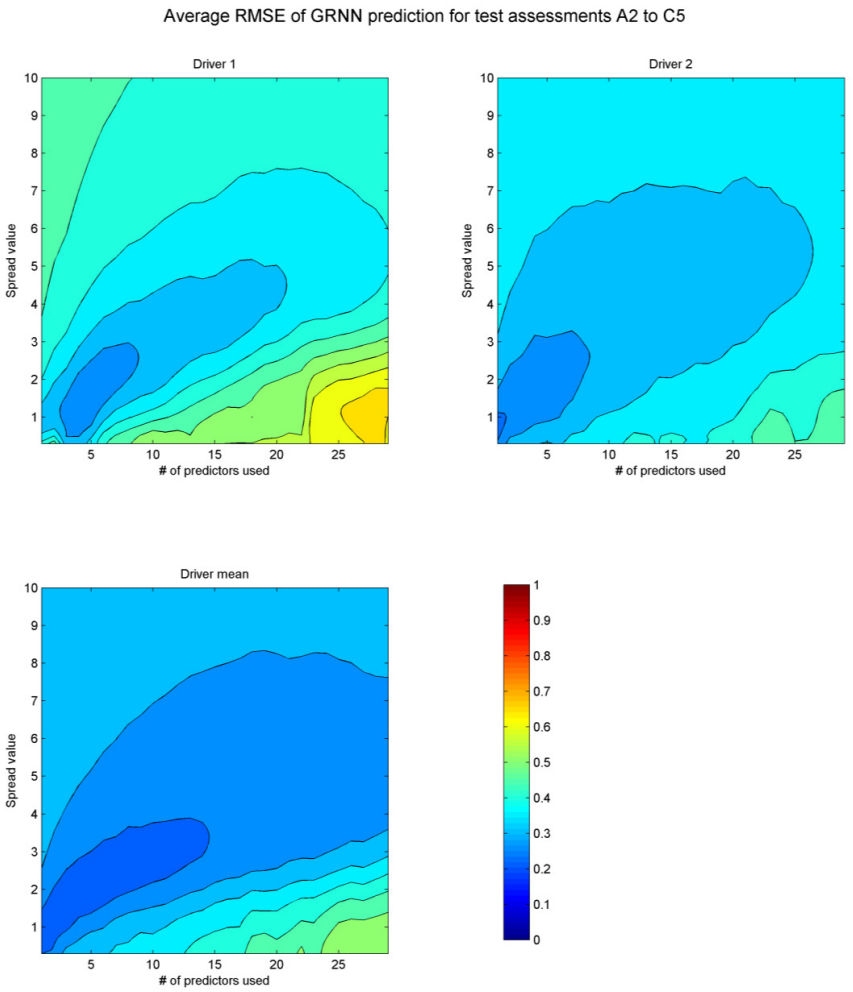


Figure 79. Result of the leave-one-out cross validation: contour plot of the average RMSE of GRNN prediction for test assessments A2 to C5 for all drivers.

The figures show that the lowest average RMSE related to best GRNN prediction performance is realized in the area with a low number of predictors used and a low spread. This best prediction area when looking at all drivers results is positioned between 1 to 8 predictors used and .3 to 3.0 as spread value with the restriction that increasing number of predictors is combined with an increasing spread.

The parameter values corresponding to the lowest RMSE for a driver are the suggested values to use when there are no targets available. To analyze the GRNN prediction performance when using these suggested parameters, these are applied for the GRNN for predicting the intermediate test assessments for all drivers. The GRNN now uses the same suggested parameter values for all test assessments for a driver. The results are shown in Fig. 80 to 82.

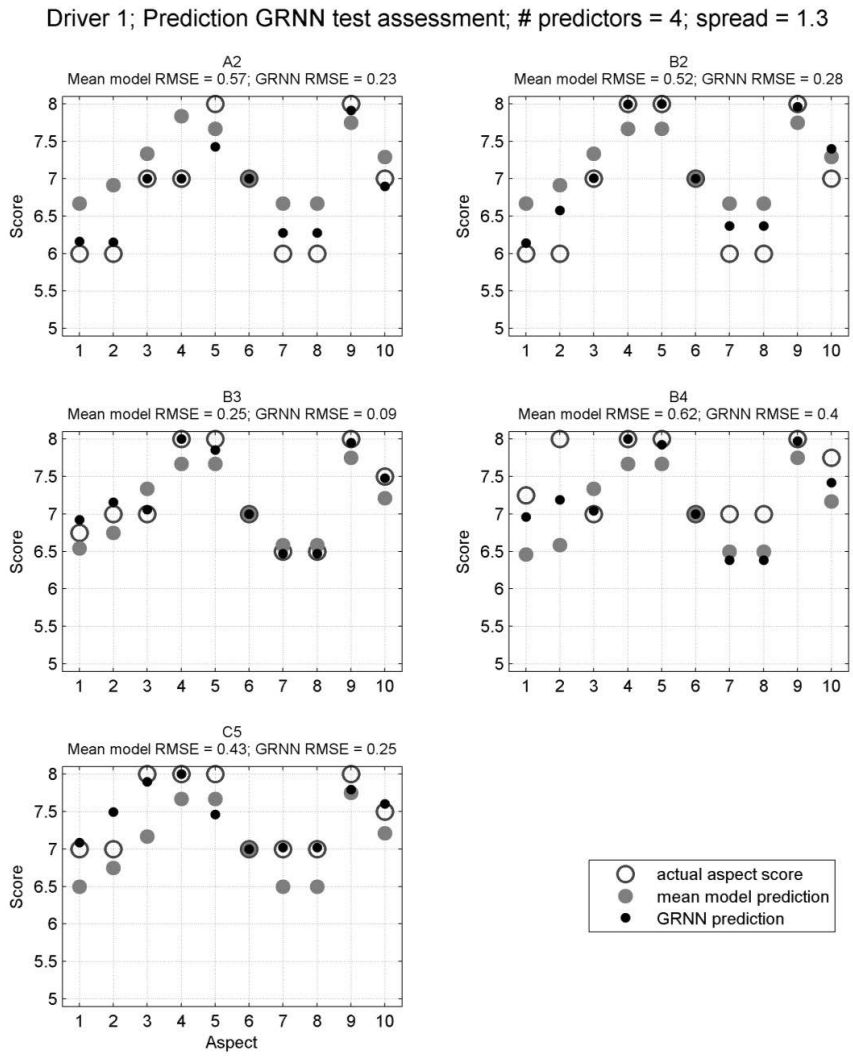


Figure 80. Driver 1 plots showing the actual, mean model and GRNN prediction aspect score for the intermediate test input assessment when using the leave-on-out cross validation suggested parameter values.

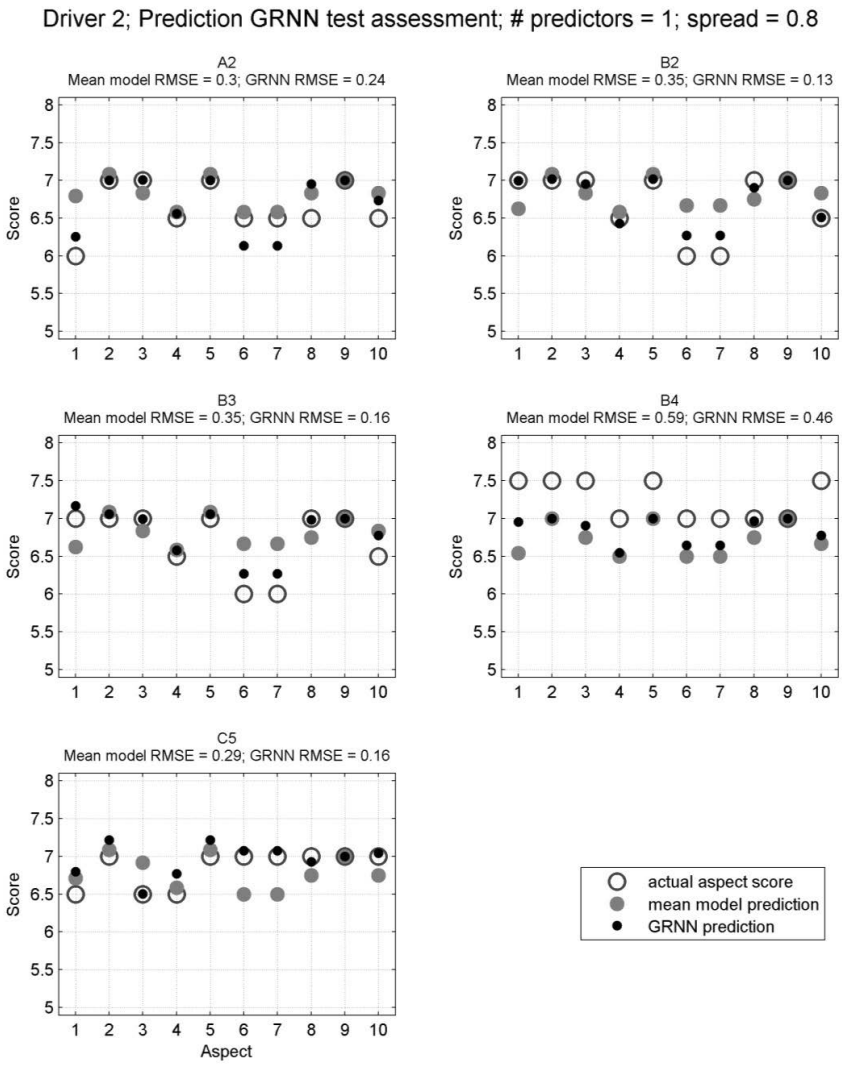


Figure 81. Driver 2 plots showing the actual, mean model and GRNN prediction aspect score for the intermediate test input assessment when using the leave-on-out cross validation suggested parameter values.

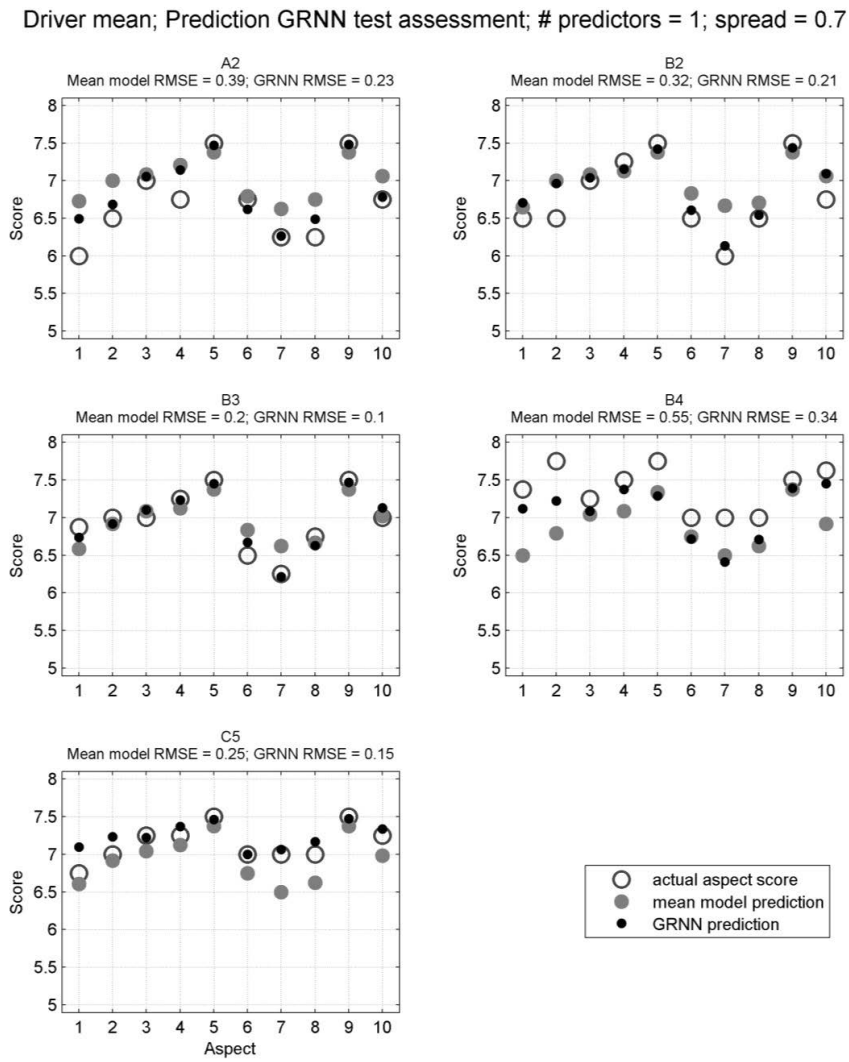


Figure 82. Driver mean plots showing the actual, mean model and GRNN prediction aspect score for the intermediate test input assessment when using the leave-on-out cross validation suggested parameter values.

These plots show good GRNN prediction performance with all RMSE values around or lower than the drivers score resolution of 0.25, except for test assessment B4. This is due to being the 'extreme assessment' for batch B and as batch C is rated relatively low, also for batch C, as explained before. All GRNN RMSE values are much lower than the mean model RMSE, which makes it a useful model for prediction.

6.5.3 Relevant Metrics

In Appendix 6 all metrics are presented per driver per aspect sorted from highest to lowest correlation coefficient value. Analyzing the metrics that have the highest correlation per aspect, and therefore will be used first as predictors, could indicate what vehicle dynamic behavior is most relevant for the driver. For example, for handling aspect 1 steering precision while cornering, the 6 metrics with highest correlation coefficient value r for this aspect are given in Table 11. The correlation coefficient value r is also shown. Low values could indicate a small variance in aspect scores, making the metrics carry less information. Aspect scores 6 of driver 1 have the same values for all tire assessments, resulting in r-values for all metrics of zero. The same holds for aspect scores 9 of driver 2. Obviously, for these constant aspect scores 'top metric' analysis is possible. Aspect 2 of driver 1 is similar, having only one aspect score 0.5 higher than all others. Although this does produce a ranking in metrics corresponding with the r-value, the information is low, reflected in the relative low r-value, with the highest value being 0.58.

Table 11. Metrics with highest correlation coefficient r-values for aspect 1, steering precision while cornering, for all drivers.

	Aspect 1 highest correlated metrics					
Driver 1	v_y-V r = 0.94	psi-I r = -0.93	M_H-I r = 0.93	$delta_H$-III r = -0.92	beta-IV r = 0.91	beta-I r = 0.90
Driver 2	a_y-I r = 0.95	a_y-II r = 0.83	v_y-V r = 0.79	a_y-III r = 0.79	r-II r = 0.63	M_H-I r = 0.62
Driver mean	v_y-V r = 0.96	v_y-IV r = 0.89	psi-I r = -0.89	a_y-I r = 0.88	M_H-I r = 0.87	beta- I/II/III/r-II r = 0.86

The metrics v_y-V and M_H-I are given with colored background to emphasize that these metrics are present in this highest correlation metric list for aspect 1 for all drivers. This can indicate that these signals and more specific, this part of the signals, is relevant for the driver to determine the handling score for that aspect. The parts belonging to the metrics v_y-V and M_H-I are shown in Fig. 83.

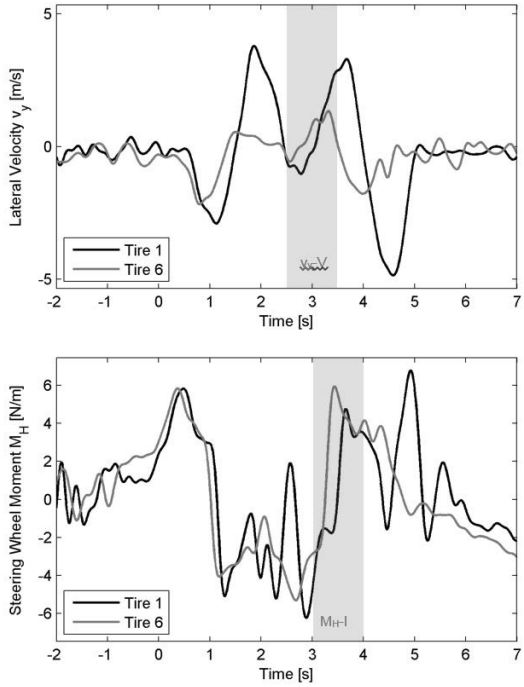


Figure 83. Signal parts belonging to metrics v_y-V (top graph) and M_H-I (bottom graph).

These signal parts indicate where the driver steers back from the second lane to the first lane. This could indicate that the driver is especially feeling the differences for the steering precision while cornering during this second part of the double lane change, where there is also influence of the first lane change vehicle behavior.

Both measures are calculated as the positive area below the signal line and have positive correlation with aspect 1. This indicates that the larger this area, the higher the score. In this case, the area is mostly determined by the delay of the signal, less delay gives a higher part of the peak in the area calculation. This is also influenced by the driven speed. This is also a metric, but this metric is not so highly correlated, being at position 12 for driver 1 and 2 and at position 11 for driver mean for the r-values for aspect 1.

For information regarding relevant metrics for more general handling behavior, so over all aspects, the results of the leave-one-out cross validation can be used. As shown in Fig. 79, the suggested number of predictors used over all drivers for best performance is from 1 to 8 predictors. If we take the suggested values that are used for Fig. 80 to 82, driver 1 uses 4 predictors and driver 2 and driver mean only 1. Table 12 shows these predictors; for readability the r-values are left out, except when $r = 0$, because then no metric is relevant.

Table 12. Metrics used as predictors for the leave-one-out cross validation GRNN's.

	driver 1 predictors				driver 2 predictor	driver mean predictor
Aspect	**#1**	**#2**	**#3**	**#4**	**#1**	**#1**
1. Steering precision while cornering	v_y-V	psi-I	M_H-I	$delta_H$-III	a_y-I	v_y-V
2. Stability while cornering (no throttle change)	$delta_H$-III	psi-I	v_y-V	beta-IV	v_y-V	v_y-V
3. Stability while cornering (throttle change)	v_x-I	M_H-II	a_y-II	beta-III	psi-III	r-II
4. Yaw overshoot	beta-II	v_y-I	a_y-III	beta-I	beta-III	v_y-V
5. Predictability	$delta_H$-I	a_y-III	a_y-V	M_H-II	v_y-V	v_y-II
6. Yaw delay	(*r*=0)	(*r*=0)	(*r*=0)	(*r*=0)	M_H-II	M_H-II
7. Steering angle	M_H-I	M_H-I	v_y-I	beta-III	M_H-I	M_H-I
8. Grip	M_H-I	M_H-I	v_y-I	beta-III	a_y-II	beta-III
9. Controllability	M_H-I	a_y-V	psi-II	v_y-I	(*r*=0)	v_y-I
10. Overall judgement	psi-I	beta-IV	v_y-I	M_H-I	beta-III	v_y-I

For driver 1, the predictors for aspect 6, yaw delay, are not mentioned, because all metrics have a correlation coefficient of zero, owing to the fact that for this aspect all assessments scores are equal. Consequently, it does not matter which predictor is used, they will all results in the correct prediction. This also accounts for driver 2, aspect 9.

The most relevant signal metrics, being used 10 times or more as predictors are shown in Table 12 with colored background and are discussed. Notable is the overall importance of the lateral velocity v_y metrics, for all drivers this metric is used 15 times. For driver mean, this is the only predictor for 6 handling aspects. Especially metric v_y-V is often used, the part representing the lateral velocity after the first lane change.

Next relevant metrics are the steering wheel moment M_H metrics. Remarkable is the dominant importance of this metric for predicting aspect 7, steering angle. Although there are metrics defined for the steering wheel angle, these metrics are not used as predictors for the steering wheel angle aspect. Also for aspect 1, steering precision while cornering, the steering wheel moment metric is more relevant than the steering wheel angle metric.

Well known of being an important variable for driver's subjective evaluation, is the vehicle slip angle β. This is confirmed by these results, the vehicle slip angle metrics are most relevant after the lateral velocity and steering wheel moment metrics. Especially for yaw overshoot and overall judgement for driver 1 and 2 and for grip for driver mean.

The metrics representing the yaw rate and the speed are only used once as predictor, which suggests that they are not relevant metrics. In contrast, they are used as first (driver 1) and only (driver mean) predictor for aspect 3, stability while cornering with throttle change. This could be explained by the fact that

this is one of the aspects that is related to driving the circles, which is not included in the metrics.

6.6 Conclusions

The GRNN for handling assessment described in this chapter has good prediction performance of driver's subjective evaluation of tire handling aspects (rated on a scale of 1 to 10) from driving on a handling track, based on objective measures taken during only the double lane change maneuver of the handling track. The GRNN can even perform well on predicting handling aspects not related to driving the double lane change.

Analysis of the GRNN performance on the leave-one-out cross validation, resulted in a best performance area. The corresponding best values for the two GRNN parameters are for the number of predictors used, between 1 and 8, and for spread, between 0.3 and 3; with the restriction that a higher number of predictors must be combined with a higher value of spread. This shows that the prediction performance of the GRNN is not so sensitive for parameter changes.

Within this best performance area, the optimal performance was reached with small values for (number of predictors used, spread value), being (4,1.3) for driver 1, (1,0.8) for driver 2 and (1,0.7) for driver mean. Given a test assessment input from the model fitting data range, but unknown to the GRNN, the GRNN is able to predict for all inputs but B4, the 10 handling aspects with a mean absolute error around or lower than 0.25, being the resolution of the drivers used for scoring the aspects. For B4 the GRNN is extrapolating for some aspects, because these do does not lie in the model fitting data range, therefore the error is slightly higher, around 0.4.

Analyzing the used predictors provides information on what vehicle dynamic behavior is relevant for the driver. For general handling behavior, regarding all aspects, this showed that metrics based on lateral velocity, steering wheel moment and vehicle slip angle were most relevant.

Next to good prediction performance, the GRNN for handling assessment has additional advantages. It works well with a limited dataset, does not need a-prior knowledge about the data or underlying model, has good prediction between data points and behaves well for extrapolation, approaching the nearest neighbor model output. Constructing the GRNN is easy with a one-pass learning phase and only two parameters, for which suggested values ranges are given and which are not so sensitive. Learning new examples during the prediction phase, with possibly new or less predictors and/or aspects than the existing examples, is easily done, incorporating new knowledge when it becomes available.

7. Driver's Handling Assessment based on Workload Measures

This chapter describes the second approach taken in this research for understanding a driver in vehicle handling as explained in section 3.1: predicting driver's subjective handling assessment based on workload measures. This shows that there are objective measures derived from heart rate, secondary task performance and steering behavior, which correlate to the driver's subjective mental (and physical) workload assessment and therefore may act as predictor for the overall judgement of tire handling assessment as shown in Fig. 84. The shaded area shows the specific implementation for the method described in this chapter of the general model given in Figure 20 of section 3.1.

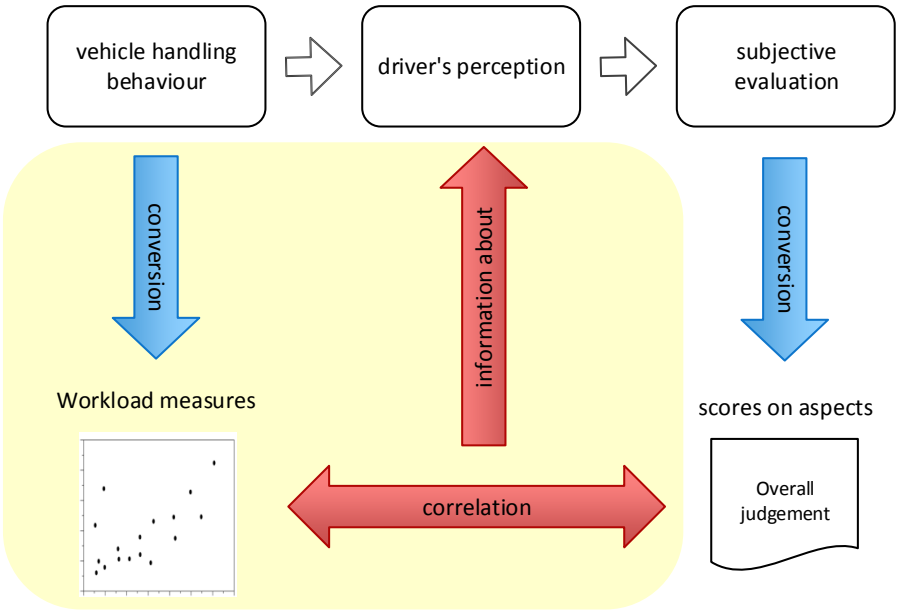


Figure 84. Correlation between workload measures and the subjective evaluation aspect 'overall judgement'.

The adapting driver theory, as described in section 2.4 and especially the relationship with mental workload (section 2.4.3 for background and references) has been the inspiration for this method. The assumption that this relationship exists, is validated in section 5.3.2. For a detailed introduction to relating objective measures, like workload measures, to subjective measures, the reader is referred to section 3.1.

To explore if this method could be interesting, a field experiment is performed, specifically designed to influence driver's mental workload by changing the task demand, *i.e.* driving slalom maneuvers and double lane change maneuvers with different tires at different speed. These two experiments are previously described in detail in respectively section 4.7.4 and 4.7.5. The methods used for these workload experiments and the derived measures are described in section 4.7.1 to 4.7.3. Assumptions made for these experiments are validated in chapter 5. In this chapter, the results of this method are described and discussed for heart rate measures, a secondary task measure and steering measures in respectively section 7.1 to 7.3. In section 7.4 conclusions are given.

7.1 Heart Rate Measures

To explore if using heart rate measures could be interesting, a field experiment with two drivers was performed. The heart rate measurements were collected during these slalom tests. From the ECG data, the heart rate and IBI data of the drivers were derived for the time domain analyses. For the frequency analyses, spectral analyses of HRV were used to calculate the LF/HF ratio as indicator for mental workload, as described in section 4.7.3. Unfortunately, for driver F many artefacts were present in the ECG-data, which could be removed by software and manually after visual inspection. This data could be used in the time domain analyses, but was not suitable for the frequency analyses, because small differences in IBI data have a large effect on the results of the spectral analyses, which makes the results unreliable. Therefore, for driver F only the time domain analyses of the IBI are presented. The maneuvers for driver C did not have these problems and could be used for frequency analyses, with an exception for one specific maneuver at high speed.

To quantify the linear relationship between a heart rate measure of the driver and his workload score, Pearson correlation coefficient r (Eq. (24)) is used.
In addition, a scatter plot of the variables is presented including the least squares linear regression line to be able to interpret r, because one or more outliers in the data can strongly influence r.

7.1.1 IBI

For both drivers for every maneuver the mean value of the IBI's is calculated, referred to as 'mean IBI'. Fig. 85 shows the scatter plot of mean IBI and mental workload scores for both drivers. A negative correlation was obtained for both drivers, with $r = -0.95$ for driver C and $r = -0.80$ for driver F.

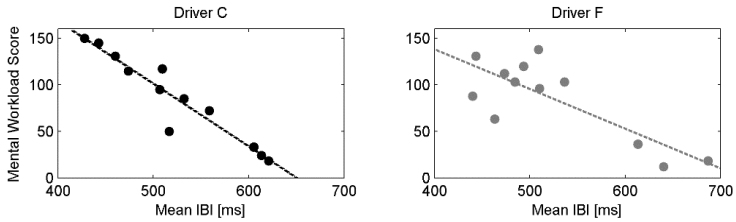


Figure 85. Correlation between mean IBI and mental workload scores for both drivers.

Both correlation coefficients indicate a strong linear relationship between mean IBI and the mental workload score. For driver C this is confirmed by the plot, as the points can be seen to lie around the least squares line. This indicates that for driver C, prediction of mental workload based on mean IBI can be made based on the regression line. It can be seen from the plot that this prediction holds especially for the low and high values of mean IBI; for the intermediate values, around 500 ms, the spread between points is larger, so the prediction will be less certain. The scatter plot for driver F show 2 'clouds' of points, with low correlation between points in the larger cloud of low mean IBI and high workload scores. The high r can be explained by the relative position of the two clouds. Often, such a spread indicates outliers, but this is not the case here, as these are valid measurements and in the expected regions. For example, for both drivers, the three high mean IBI values combined with low workload scores belong to the low speed demands of 30 and 45 km/h and therefore were expected to have low workload. This plot indicates that for driver F his mean IBI can only be used to categorize his mental workload in regions of low (score < 50) and high (score > 50) mental workload.

The disadvantage of using mean IBI for indicating mental workload is that it can also be influenced by other factors, especially physical workload (see section 2.4.3 to 2.4.5). As explained in section 4.7.2, for the field experiments, the driver's mental workload was increased by increasing the primary task demand, to match tire testing practice. Influencing the primary task demand by changing cone spacing, speed and/or tire pressure inevitably influences the physical workload of the driver, while this involves more irregular and faster steering. This is confirmed by Fig. 86, where the scatter plot of mental workload scores and physical workload scores for both drivers are shown. A positive correlation was obtained for both drivers, with $r = 0.95$ and $r = 0.97$ for drivers C and F, respectively.

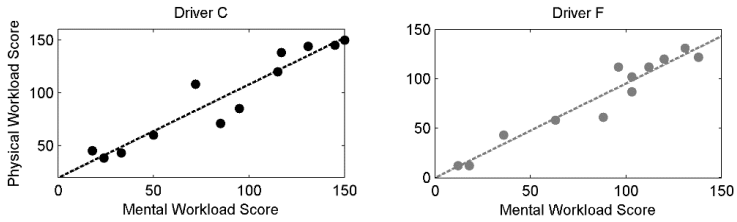


Figure 86. Correlation between mean mental and physical workload scores for both drivers.

This shows a strong linear relationship between mental and physical workload and makes it not possible to separate their influence on the heart rate measures, because both have the same underlying independent variable: the task demand of driving the slalom.

Based on this strong correlation between mental and physical workload for driving in this experiment, it is expected that also the physical workload is correlated to the mean IBI. Fig. 87 shows this scatter plot of mean IBI and the physical workload scores for both drivers. A negative correlation was obtained for both drivers, with $r = -0.88$ and $r = -0.73$ for drivers C and F, respectively.

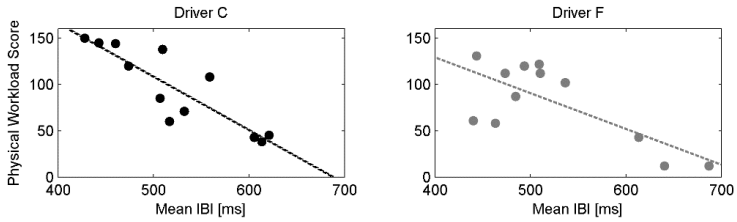


Figure 87. Correlation between mean IBI and physical workload scores for both drivers.

These results confirm that for this experiment, mean IBI is also strongly correlated with physical workload, but the correlation is less strong than found for mental workload.

7.1.2 HRV

As explained in section 2.4.5, the power in the band around the 0.10 Hz component of the HRV is related to mental workload, especially for task related effort, with the advantage that other influences, like physical workload, are small. Therefore, it is of interest to see if this measure can be useful for tire testing where physical and mental workload is strongly related. For the HRV the ratio of the power in the low frequency (LF) range between 0.04-0.15 Hz and the power of the high frequency (HF) range between 0.15-0.4 Hz is taken as measure. An increase of the LF component and decrease of the HF-component can indicate mental workload (Guger et al., 2004). Fig. 88 shows an example

the power spectral density plot of HRV data, with the LF and HF range indicated.

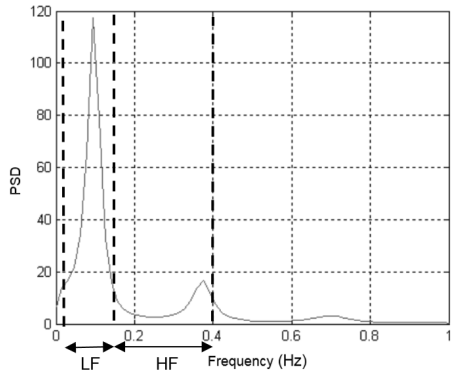


Figure 88. Example of a Power Spectral Density plot of HRV.

As indicated previously, only for driver C frequency analysis of the heart rate measurements was possible. For 11 maneuvers of this driver, the LF/HF ratio was calculated and related to his mental workload score. Fig. 89 shows the scatter plot of the LF/HF ratio and the mental and physical workload scores for driver C with respectively, $r = 0.92$ and $r = 0.90$.

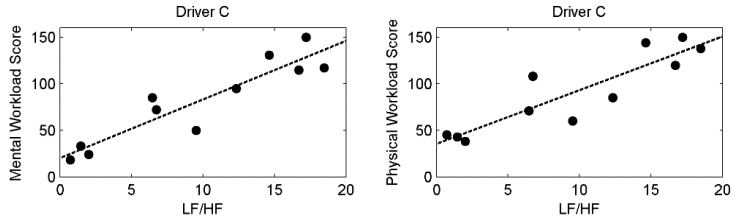


Figure 89. Correlation between LF/HF ratio of the HRV and mental and physical workload scores for driver C.

Both correlation coefficients indicate a strong linear relationship between LF/HF and mental and physical workload, a little stronger for mental workload.

7.1.3 Summary

Table 13 shows a summary of the correlation coefficients found between both mental and physical workload scores and mean IBI for both drivers and between these workload scores and the LF/HF ratio for driver C.

Table 13. Correlation coefficients r.

	Driver C Workload Score		Driver F Workload Score	
	Mental	**Physical**	**Mental**	**Physical**
Mean IBI [ms]	−0.95	−0.88	−0.80	−0.73
LF/HF [-]	0.92	0.90		

All the workload scores show a strong linear correlation with the measure mean IBI. For driver C this measure can be used to predict his workload, because the plot also shows a small spread along the regression line. For driver F, prediction of workload by mean IBI is limited to categorizing workload in low and high regions, due to the larger spread of points in the scatter plot. IBI is known to be influences by other factors than mental workload, which are driver dependent, like physical condition and emotions (Waard, de, 1996), which could explain the differences in spread in IBI values for the drivers. During tire testing, and confirmed in this experiment, there is a strong correlation between mental and physical workload. A measure known to be less influenced by other factors, is a HRV measure, the ratio LF/HF. The results confirmed that this measure also correlated strongly with mental workload, but as indicated before, physical workload could not be separated in this experiment. The derivation of this HRV measure by spectral analyses of IBI does requires more data processing and does not allow too many artefacts in the ECG data, which is more difficult to prevent in applied automotive settings than in laboratory setting.

7.2 Secondary Task Measure

During the double lane change workload experiment, a secondary reaction task was implemented, as a measure for mental workload, as described in section 4.7.3. Decreasing performance on the secondary task, being a longer reaction time, should indicate increasing driver's mental workload for the primary task, being driving the double lane change with the predefined speed without hitting cones. Fig. 90 shows the results for the mean value of the secondary reaction time, grouped by drivers group, tire and speed.

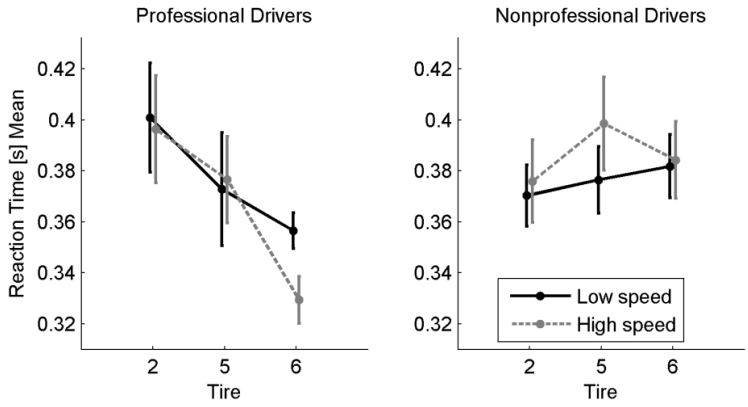


Figure 90. Mean values of the secondary reaction time task, grouped by driver group, tire and speed.

Against expectations, no clear speed dependency for the secondary reaction time can be seen. In contrast, tire dependency can be seen for the secondary reaction time for the professional drivers, although the error bars have quite an overlap. No clear speed or tire dependency is seen for the nonprofessional drivers, having a slightly increasing trend for the better handling tires.

The professional drivers' results suggest that workload induced by speed is different from workload induced by different handling tires, the latter being related with this measure.

7.3 Steering Measures

In the experiments, increasing the primary task demand was used to increase the driver's mental workload. This also increased the driver's physical workload as explained before. As steering measures are strictly taken, physical workload measures, these are often used as mental workload measures (see section 2.4.3), due to the close connection of mental and physical workload for driving. In this section, two steering measures derived from the experiments are analyzed.

7.3.1 Steering Wheel Angle Peak

An easy to obtain measure from the steering wheel angle data is the steering wheel angle peak value, from the double lane change workload experiment. This measure is derived from the first lane change of the maneuver, as this is done from the steady state condition of driving to the double lane change with constant speed. The difference in speed and tire is therefore more dominant than for following steering actions, when the vehicle is not in steady state conditions. In this case, the steering wheel angle is also influenced by the previous vehicle state. Fig. 91 shows the mean values of the first steering wheel angle peak, grouped by drivers group, tire and speed.

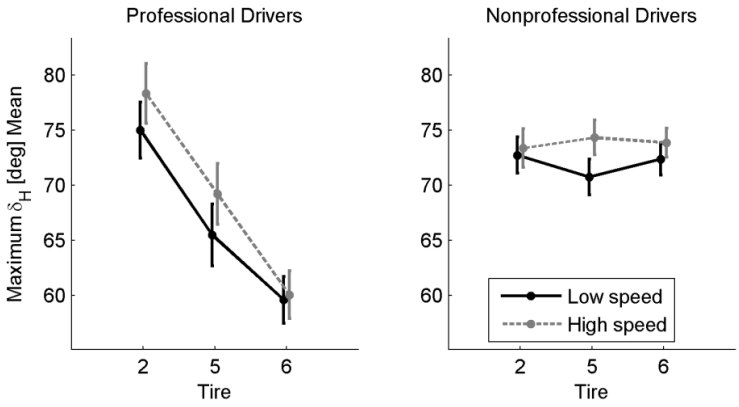


Figure 91. Mean values of maximum value of the first peak of the steering wheel angle δ_H, grouped by driver group, tire and speed.

This figure shows that the mean maximum steering wheel angles of the first peak are higher for high speed than for low speed for both drivers groups and all tires, although not so much for tire 6 for all drivers and tire 2 for the nonprofessional drivers.

This measure is clearly tire dependent for the professional drivers, having no overlap in the SEM error bars. The value decreases with better handling tires, indicating that better handling tires need less steering wheel angle input for the professional drivers. No clear tire dependency of the steering wheel angle peak value can be seen for the nonprofessional drivers. This can indicate that the professional drivers are better able to explore the better handling of the tires.

As with the secondary task measure, the professional drivers' results suggest workload differences for speed and tires; this measure indicating tire workload differences.

7.3.2 High Frequency Area

From the slalom experiment, the power spectral density (PSD) plots of the steering wheel angle are derived. Fig. 92 shows these plots of driver C for two distinctive slaloms, one at 30 km/h with constant cone spacing of 16 m and one at maximum speed with varying cones spacing of 16/20 m. The peak at the lowest frequency is explained by driving on the circle of the slalom, the highest peak comes from driving the slalom. It can be seen that for the extreme slalom the highest peak is wider and contains more high frequencies. From the PSD plots of this experiment, these 'slalom-related frequencies' all lie in the area below 2 Hz. Higher frequencies are also present, but these can be related to fast corrective steering actions performed by the driver. These high frequencies are more visible if the PSD's are plotted in decibels, as is done in Fig. 93.

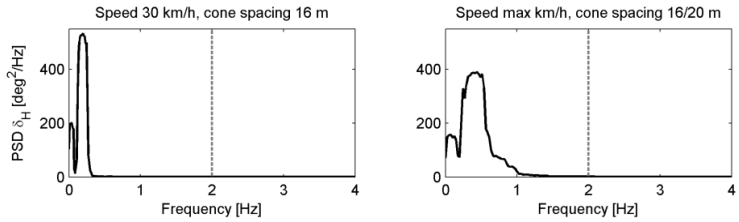


Figure 92. PSD plots of steering wheel angle δ_H for driver C for the slalom at 30 km/h with 16 m cone spacing and at maximum speed of 61.5 km/h with cones spacing of 16/20 m.

Fig. 93 shows that for the more extreme slalom, more power is present in the high frequency area above 2 Hz compared to the low frequency. The more dominant this high frequency area, the more excessive steering is applied by the driver to control the vehicle.

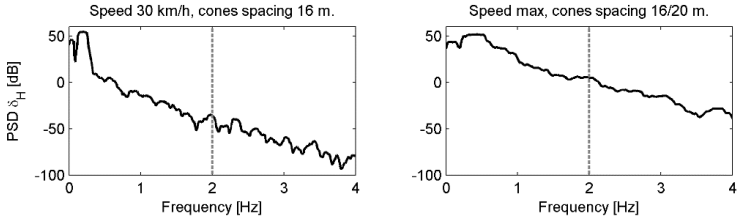


Figure 93. PSD plots of steering wheel angle δ_H for driver C for the slalom at 30 km/h with 16 m cone spacing and at maximum speed with cones spacing of 16/20 m in decibels.

To relate the high frequency area to the low frequency area, the High Frequency Area (HFA) measure is defined as the ratio of both areas, where the high frequency area is limited to 4 Hz.

Fig. 94 shows the scatter plot of HFA and mental workload scores for both drivers. A positive correlation was obtained for both drivers, with $r = 0.89$ for driver C and $r = 0.84$ for driver F.

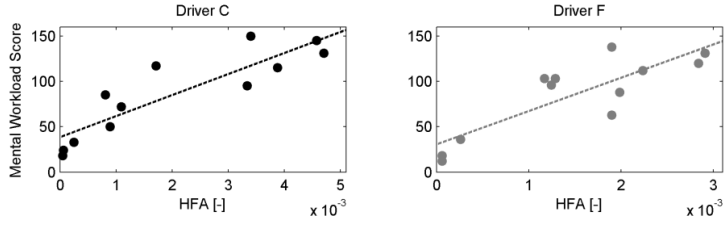


Figure 94. Correlation between HFA and mental workload scores for both drivers.

These results confirm that more high frequency steering is accompanied with a higher mental workload of the driver.

To compare the HFA measures with physical workload, Fig. 95 shows the scatter plot of HFA and physical workload scores for both drivers. A positive

correlation was obtained for both drivers, with $r = 0.83$ for driver C and $r = 0.80$ for driver F. As expected, this measure shows also a strong correlation with physical workload, although the correlation is a little less strong as with mental workload.

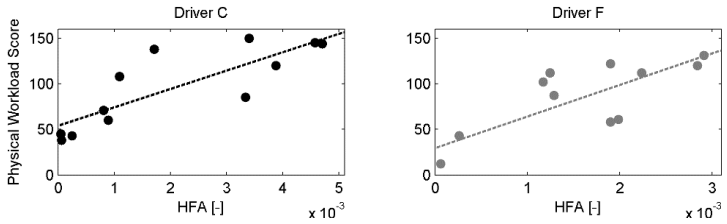


Figure 95. Correlation between HFA and physical workload scores for both drivers.

For the double lane change experiment, this measure is also calculated. Fig. 96 shows the PSD's of the steering wheel angle for two distinctive double lane changes for professional driver 1, one with tire 6 at low speed and one with tire 1 at high speed.

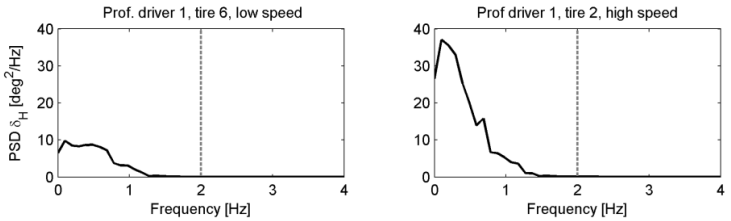


Figure 96. PSD plots of steering wheel angle δ_H for professional driver 1 for tire 6 at low speed and tire 2 at high speed.

This figure shows that the high frequency area is above the frequency peak due to driving the double lane change. The high frequency content can therefore be related to corrective high frequent steering actions of the driver. To visualize the high frequency content, Fig. 97 shows the same double lane changes in decibels.

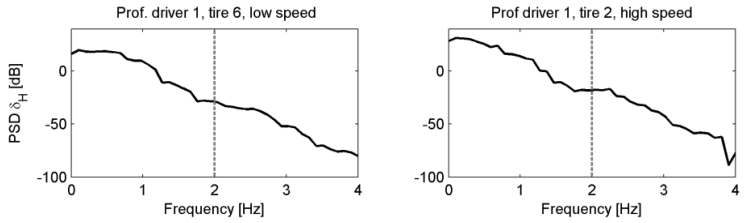


Figure 97. PSD plots of steering wheel angle δ_H for professional driver 1 for tire 6 at low speed and tire 2 at high speed in decibels.

From these PSD's the HFA measures are calculated as defined previously. Fig. 98 shows the mean values of the HFA measure, grouped by drivers group, tire and speed.

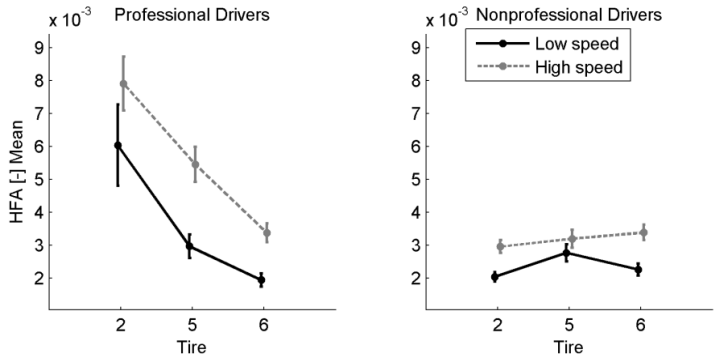


Figure 98. Mean values of HFA grouped by drivers group, tire and speed.

This figure shows that the mean HFA values are speed dependent, being higher for high speed than for low speed for both drivers groups and all tires. This indicates that for higher speed, the drivers have to perform more high frequent steering actions.

The mean HFA values are clearly tire dependent for the professional drivers, having no overlap in the SEM error bars. The HFA value decreases with better handling tires, indicating that better handling tires need less high frequent steering actions. No clear tire dependency of HFA can be seen for the nonprofessional drivers.

7.4 Conclusions

In this chapter, several objective measures are derived and their correlation to driver's experienced mental workload assessment is analyzed to see if these measures could be used as predictors for tire handling assessment.

Measures derived from heart rate data for two drivers are mean IBI and for one driver HRV. Both measures show high correlation with driver's mental workload in handling maneuvers, which makes these measures promising as tire handling predictors. As expected for handling maneuvers, which are steering intensive, the results showed that driver's mental and physical workload are highly correlated. Both measures show stronger correlation for mental than for physical workload, especially HRV. This confirms the results found in literature that this measure is suitable to filter out physical workload from mental workload for driving tasks. The results of this exploratory research for using heart rate measures to indicate driver's mental workload are promising, but require further research with a larger group of participants to make the results broader applicable. In addition, the collection of heart rate data

as used in this research was sensitive to noise. For broader application, easier and less noise sensitive ways of measuring driver's heart rate will be crucial.

The secondary task reaction time even as the steering wheel angle peak showed that this measure is able to distinguish between tires for the professional drivers. For the nonprofessional drivers this was not clear.

Steering measure HFA showed a clear speed and tire dependency for the professional drivers, for the nonprofessional drivers this was seen for the slalom experiment, but not for the double lane change.

From these results, the HFA seems to be the best measure to indicate drivers' workload differences due to speed and handling performance of tires for professional drivers. For nonprofessional drivers, this can also be seen, but this depends on the handling maneuver. In addition, HFA is a relative easy measure to obtain; in contrast to the secondary task measure and the promising, but more difficult to obtain, measures from heart rate.

8. Driver's Handling Assessment based on Driver Model Parameters

This chapter describes the third method taken in this research for understanding a driver in vehicle handling: predicting driver's subjective handling assessment based on driver model parameters. The background and reason why this method was chosen was explained in section 3.2. This driver model method is mainly related to the question *how* the driver perceives the tire handling behavior. Driver aspects that are considered are mental workload, physical workload, adaptation and anticipation to changing driving circumstances and vehicle behavior. For this method, the driver is modelled in closed-loop with the vehicle, as shown in Fig. 99.

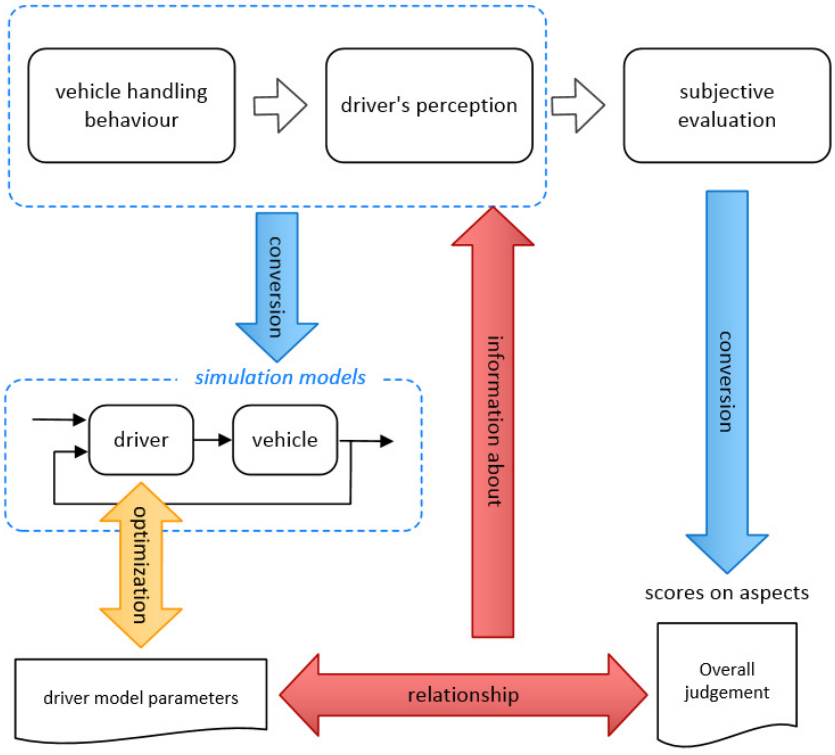


Figure 99. Driver model method: relating driver model parameters to the overall judgment of the tires.

The driver is modelled with explicit parameters, which are related to driver's characteristics like preview time, steering gain, delay, lag and lead behavior. Different tires used in testing are simulated and the driver model parameters are optimized to fit the simulation results to the test results. Analysis of the driver model parameters could provide information about the aforementioned driver aspects, mental workload being an important one, as explained in the adapting driver theory in section 2.4. Relating these outcomes to the subjective assessment of the tires, in this case on the aspect 'overall judgement', could reveal relevant aspects in the driver's perception of the tire handling behavior. This can provide 'how'-information on the driver's perception.

To verify this driver model method, a field experiment is performed, specifically designed to influence driver's mental workload by changing the task demand, *i.e.* driving double lane change maneuvers with different tires at different speed. This experiment is described in detail in section 4.7.5, with the general methods used for these workload experiments described in section 4.7.1 to 4.7.3. The validation of the assumptions made for this research are described in section 5.3.

This chapter starts with the survey, selection and the preparatory application of the driver model for this method in section 8.1. In the next section, the method for the model parameter estimation is described. The results are given in section 8.3 and the conclusions are given in the last section 8.4.

8.1 Driver Model Survey, Selection and Application

To select a model for this driver model method, a literature survey was performed as part of a preliminary study of tire test driver models (Arts, 2007).

8.1.1 Driver Model Requirements

The driver model applied in this method is used for explaining driver behavior; therefore, the model parameters should represent relevant realistic driver behavior. This resulted in certain requirements for this driver model. MacAdam gives a thorough overview and description of understanding and modeling the human driver (Macadam, 2003). For modeling the human driver, he defines certain human properties and characteristics, essential in a model for representing realistic human driver behavior:

- Time delay
- Preview, the driver must look ahead, so he can anticipate with his control
- Adaptation, so the driver can adapt to different vehicle behavior
- Behavior according to the McRuer crossover model
- Estimation of vehicle response to control input

In addition to these essentials, for this research, the driver model should include the following characteristics:

- Lateral control, the driver steering task is essential for tire handling. As the tire handling tests done for this research are performed with constant speed, longitudinal control is not necessary.
- Preview tracking, the tire handling tests done for this research required the drivers to follow a certain path, and preview was already part of the aforementioned essentials.
- Explicit parameters for the driver, which can be related to driver characteristics.

8.1.2 Selected Driver Model

Given these driver model requirements, a selection of driver model literature was made. This selection can be categorized in driver models with combined lateral and longitudinal control (Edelmann, Plöchl, & Lugner, 2007; Prokop, 2001; Sharp, Casanova, & Symonds, 2000), additional background on human preview behavior (Cole, Pick, & Odhams, 2006; Donges, 1978; Salvucci & Gray, 2004; Sharp et al., 2000), summary of different driver models (Macadam, 2003; Plöchl & Edelmann, 2007; Reid, 1983), explicit path following models (Sharp et al., 2000; Toffin, Reymond, Kemeny, & Droulez, 2007), adaptive preview driver models (Cole et al., 2006; Edelmann et al., 2007; Prokop, 2001; Toffin et al., 2007) and quasi-linear driver models, with optimization (Horiuchi & Sunada, 1998; Horiuchi & Yuhara, 1996; Pauwelussen & Pauwelussen, 2004; Yuhara & Tajima, 2000) and without optimization (Donges, 1978; Reid, Solowka, & Billing, 1981; Reid, 1983; Salvucci & Gray, 2004; Sharp et al., 2000).

Based on the driver model requirements and the selected literature, the preview path driver model as shown in Fig. 100 was defined for this research. The driver model is divided into two parts: a perception part and a control part, which are explained in detail in the next two paragraphs. The general flow of information for this driver model is as follows. The perception part of the driver perceives the input, *i.e.* the target path he wants to follow and feedback of the vehicle state, like actual vehicle position, orientation and velocity. The perception part evaluates this input and will result in an error as output, called the preview error $e_{preview}$, if the driver expects the vehicle to deviate from the target path at a certain preview point ahead, if no steering control is given. This error is the input for the driver control part, which results in a steering wheel control action to the vehicle with the objective to reduce the actual error when the vehicle will arrive at this preview point.

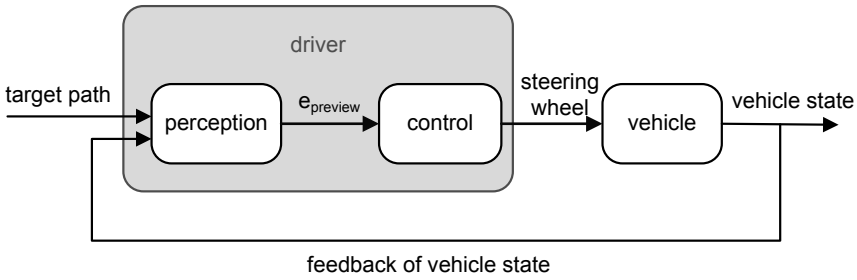


Figure 100. Path preview error driver model.

Driver Model Perception

For a real driver, perception is multimodal; information received by the visual, haptic, vestibular and auditory system is combined to perceive the state of the driver-vehicle system. For this driver model, like for most driver models, only visual information is taken as input, as visual information is essential for a driver and by far the most dominant information in driving (Macadam, 2003). Fig. 101 explains the perception part of the driver model. The input for the driver is a predefined target path he wants to follow. In the figure, the vehicle is not positioned at the target path. The difference between the actual vehicle position and the target path vehicle position with the same X-value in the (X, Y) global axis system, is the actual path error e_{actual}. This assumes the target path orientation and the basic vehicle movement to be along the X-axis with small yaw angle values, like when driving a double lane change. The driver is assumed to look straight ahead of the vehicle at a preview point a preview distance ahead. This preview distance $d_{preview}$ is one of the driver model parameters. The difference between this preview point and the point on the target path with the same global X-position, the target preview point, gives the preview error $e_{preview}$. Based on this preview error the driver will give a steering control to the vehicle by setting the steering wheel at a certain angle to get the vehicle back on the target path, in this case the driver will therefore steer to the right.

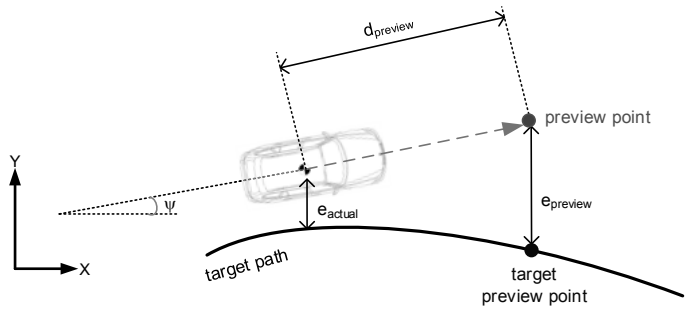


Figure 101. Perception part of the preview error driver model.

For this model, the yaw angle ψ is taken into account, which defines the orientation of the vehicle in the global axis system (X, Y). In the previous study (Monsma & Arts, 2008) also the yaw velocity was taken into account, to include more realistic driver behavior. For example, if the vehicle in the figure is already turning to the right, having negative yaw velocity, the driver will include this information in his steering control. In the extreme case, if the vehicle is turning very fast to the right, this may even result in a driver control steering wheel angle to the left. In the current study, the yaw velocity is not directly included, as it would double the number of driver control parameters, but indirectly included by the lead term in the driver control part (explained in the next paragraph). This term concerns the change in $e_{preview}$, which is strongly related to this yaw velocity (next to the change in the target path). Taking the same example, the vehicle turning to the right will make the preview error smaller, which will be taken into account in the driver steering control model by the lead term.

Driver Model Control

The control part of the driver model is the quasi-linear model, described in section 2.4.1 and given in Eq. (9). This driver model has a total of six driver parameters: one from the perception part, the preview distance $d_{preview}$ and five from the control part. The two driver limitations parameters are taken as constants in this research, because the variation in these parameters are in general limited as explained in section 2.4.1, even more because of the homogeneous group of drivers that took part in this research (group of concentrated and highly skilled drivers performing repeated testing).

8.1.3 Driver Model Application

The driver model was first explored in a simplified version in a simulation study. The objective of this study was to evaluate if changing tire characteristics could be related to changing driver model parameters (Monsma & Arts, 2008). The results showed that the driver model parameters were able to adapt to changing tire characteristics, based on a balance between mental workload and handling performance.

The next step in this research was a study to see whether this same objective could be achieved using real test data (Monsma & Oort, van, 2010). Test data of the tests described in section 4.6 were used for three different tires that were tested subjectively in one batch. The model parameters that could adapt to the different tires, were the preview time $t_{preview}$ (being equal to the preview distance $d_{preview}$ divided by the (constant) longitudinal velocity), the gain K and lead term time constant τ_{lead}. The results showed an almost constant $t_{preview}$ of 0.8 s for both drivers and all tires, which was comparable with other values found for (double) lane changes (Abe, 2009; Macadam, 2003; Tseng et al., 2005), although lower values of 0.3 s were also found (Droogendijk, 2010) and for more than one preview point, values between 0 and 1.0 s are used . This implies that the preview distance increases with increasing constant speed. The values for K and τ_{lead} were both driver and tire dependent, but only τ_{lead} showed correlation with the subjective assessment: the higher the subjective

score of the tire, the lower τ_{lead}. This parameter is related to anticipation of the driver and requires more driver effort to apply, as explained in section 2.4.1 and 2.4.2. This could therefore be a useful parameter related to mental workload and subjective evaluation of the tires.

8.2 Model Parameter Definition and Estimation

For driver model parameter identification, the driver model can be simulated in an open-loop or closed-loop configuration. Previous research (Oort, van, 2011) showed that open-loop identification of driver model parameters, taking only the driver model input and output into account, could result in driver model parameter values that, when applied in the closed-loop driver-vehicle model, could give unstable driver-vehicle performance, therefore not reflecting the real driver behavior. In addition, this method must be applicable for virtual testing, which implies the use of closed-loop simulations of the vehicle driver system. Therefore, for this research the driver model is simulated in closed-loop. For closed-loop simulation, a vehicle model must be included, which is explained next.

8.2.1 Vehicle Model

The test vehicle was modelled with a nonlinear single track vehicle model with nonlinear axle characteristics, with adjustable parameters to account for effects like load transfer and changing test conditions like weather, temperature and road humidity (Genta, 1997; Pacejka, 2012; Pauwelussen, 2014), that was previously developed for this research (Oort, van, 2012). To let the vehicle model represent the actual vehicle behavior as much as possible, the vehicle model was fitted in open-loop on the test data. For every tire set, the identified (see section 4.2) Magic Formula nonlinear tire models including relaxation behavior were included in the vehicle model and were each fitted to the test data by using lateral force scaling factors for every axle. For all double lane changes these two scaling factors, together with the yaw rotational inertia of the vehicle were optimized in least squares sense by fitting the model outputs on the measured outputs for the yaw velocity and the lateral velocity. This allowed the vehicle model to adapt to the aforementioned effects. To avoid model overfitting, the average values of the optimized parameter values of one test cluster (a combination of driver-tire performing consecutive double lane changes) are taken as parameter values in the vehicle model for that test cluster for the closed-loop simulations to identify the driver model parameters (as explained in the next section). Fig. 102 shows the mean values of the vehicle model input and output data and measured data for the maneuvers driven with the (extreme) combination of tire 6 driven by the professional drivers for the high speed demand. This figure shows that the vehicle model is able to represents the actual vehicle behavior. All other combinations show similar behavior for the model compared with the measured values.

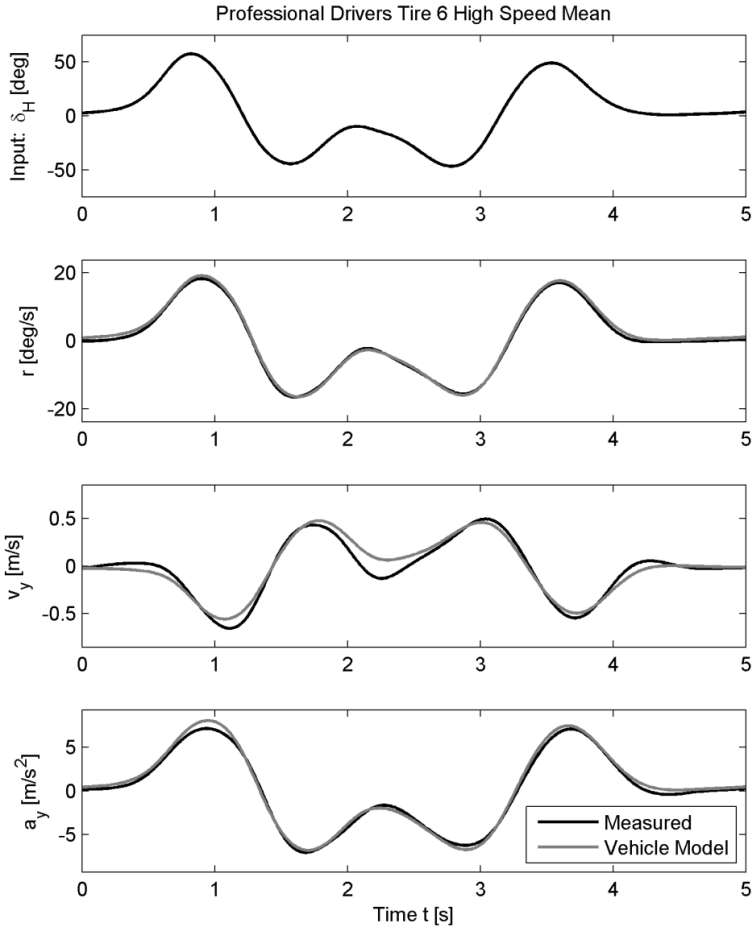


Figure 102. Mean values of the open-loop vehicle model input and output data and measured data for the maneuvers driven with tire 6 by the professional drivers group for the high speed demand.

8.2.2 Driver models

The selected driver model described in section 8.1.2 and shown in Fig. 100 is applied in three different implementations, depending on the way the target path is determined, the number of parameters used and if only feedback or also feedforward information was used. These three different implementations are referred to as driver models A, B and C. Optimizing many parameters at the same time, especially if they are not independent, can give less reliable results when the parameter values are used to explain real phenomena. An example for driver models are the parameters gain and preview distance, when they are

optimized together, it was found that these parameters are strongly related: shorter preview distance results in higher gain and vice versa (Pauwelussen, 2012). This means that (very) different combinations of both parameters, can give the same result. As long as this can be justified with actual driver behavior, as in this example, this is not a problem, but if there are many mutually dependent parameters, the results can be hard to explain. Hence, using as few parameters as possible, but still enough to be able to explain the realistic driver behavior, is good modeling practice. This practice is followed for the three implementations of the driver model, where for each implementation more driver parameters are added with the intention to include more realistic driver behavior, but being accompanied with the risk of adding too many parameters, which can undermine the results.

Not all driver model parameters as given in Eq. (9) are adaptable. Two represent human limitations and cannot, or only marginally, be adapted by the driver, *i.e.*, the driver delay time τ and the neuromuscular lag time τ_N. Both values are set to the average value between limits found in literature (see section 2.4.1), giving 0.15 s for τ and 0.06 s for τ_N. The lag time constant τ_{lag} is also taken as a constant value. As the steering task of the vehicle for path following is a dominant second order control, this lag time constant must be small as explained in section 2.4.2 and is of less importance than the lead time constant in relationship to mental workload, or the gain in relationship to physical workload. In combination with the aforementioned risk of optimizing mutual dependent parameters, this lag time parameter is therefore set to a constant value of 0.08 s, which gives no significant integrator behavior in the region of the crossover frequency and (after multiplication with the neuromuscular lag) gives a dominant time constant of 0.14, which is a value typically found in driver models for overall lag time. This resulted in the following implementation for the control part of the selected driver model for all three driver models:

$$H(s) = \left(\frac{e^{-0.15s}}{0.06s + 1}\right) \cdot \left(\frac{\mathrm{K}(\tau_{\mathrm{lead}} s + 1)}{0.08s + 1}\right) \tag{29}$$

with two adaptable driver control parameters, gain K and lead time constant τ_{lead}.

Driver Model A

Fig. 103 shows the implementation of driver model A. It has three adaptable model parameters: one driver perception parameter the preview distance $d_{preview}$ and the two driver control parameters the gain K and the lead term time constant τ_{lead}.

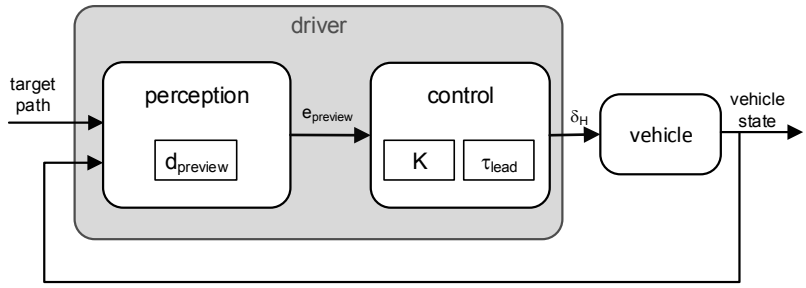


Figure 103. Driver model A with three adaptable driver parameters.

The target path is the path the driver wants to follow. For this driver model, a target path is determined for every driver as a continuous path, being the average of all driven double lane changes of that driver. Because the double lane change is bounded by cones and driven with high speed, the differences between driver's targets paths are small. Fig. 104 shows the target paths of the drivers. Note that the lane offsets for both driver groups were set at different values, to keep all drivers in the task related effort region *A*3 (see section 4.7.5).

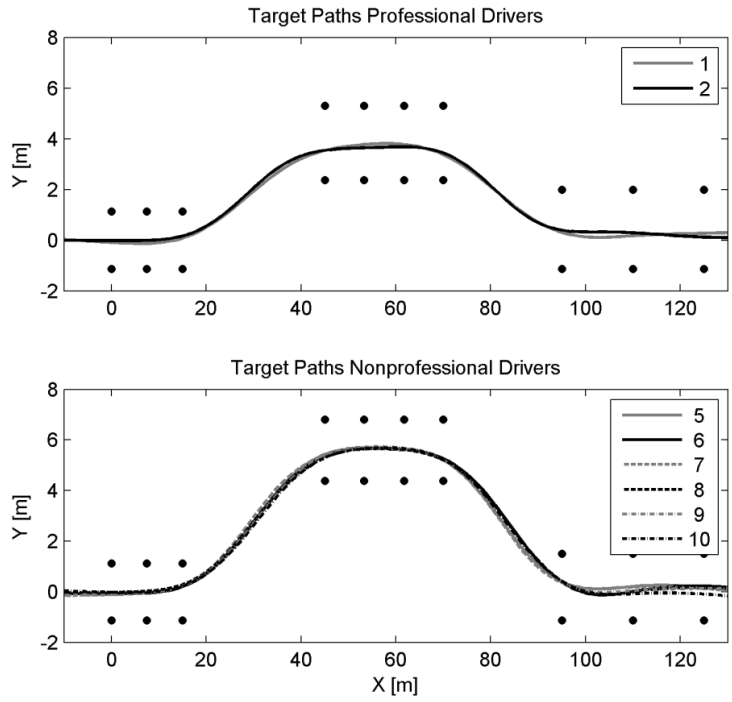


Figure 104. Continuous target paths for the individual drivers.

Although the driver looks ahead a distance $d_{preview}$, this is not regarded as using feedforward information, because feedback information is needed to predict the preview error on which the driver's steering control is based. Using feedforward information would result in driver's steering control action based only on the input target path information. Therefore, for driver model A, no feedforward information is used.

Driver Model B

The difference between driver model A and driver model B is the implementation and perception of the target path. For driver model A, a smooth target path is assumed, similar to the average driven path of that driver, which the driver looks ahead with a distance $d_{preview}$. Based on discussions with drivers and observing driver behavior during the lane changes, an alternative target path and perception part was implemented. During driving a double lane change, a driver has only one goal: keep the vehicle between the cones. The target path is therefore defined as the middle position between the cones of every lane. This results in a discrete path with only two Y-positions: the middle of the first and last lane and the middle of the second lane. The point where the driver focuses on the next lane target path, before actually reaching it, is implemented in two driver model parameters: d_{prev1} and d_{prev2}. Allowing the parameters to adapt to the differences in dimensions between the two lane changes and to the differences in vehicle state. Driving the first lane change starts with a steady state situation, whereas driving the second lane change starts with a transient situation due to the first lane change. Fig. 105 shows this discrete target path and the previewed target path, which is determined by d_{prev1} and d_{prev2}. The driver, while driving in the first lane, will start looking at the second lane target path at $X = X_{prev1}$, which is d_{prev1} meters before he will actually reach the first cone X-position of the second lane at X = 45 m. As a result, the driver will start steering to the left to reach this lane. Driving in this second lane, at $X = X_{prev2}$, which is d_{prev2} meters before reaching the first cone X-position of the last lane at X = 95 m, he shifts his focus to the last lane target path and starts steering to the right.

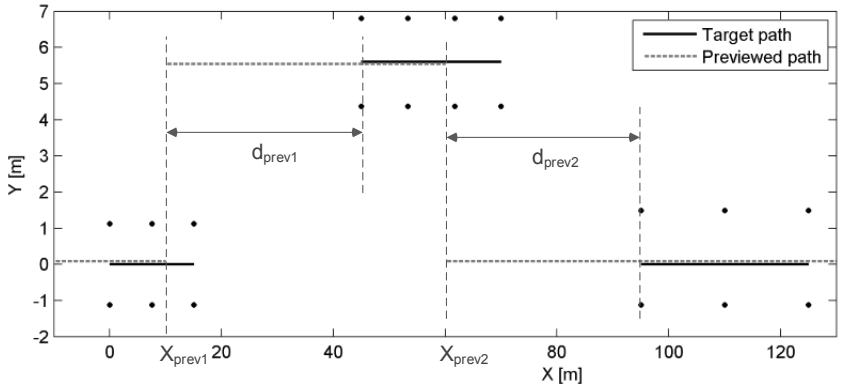


Figure 105. Discrete target path.

Although this discrete target path requires two preview distance driver parameters to be fitted instead of one, the definition of the target path is much more straight forward than for the continuous target path of driver model A. Especially when no driven paths are available, as would be the case with virtual testing.

The control part of driver model B is equal to driver model A, with two adaptable control parameters gain K and lead term time constant τ_{lead}. Fig. 106 shows the implementation of driver model B with the four adaptable driver model parameters.

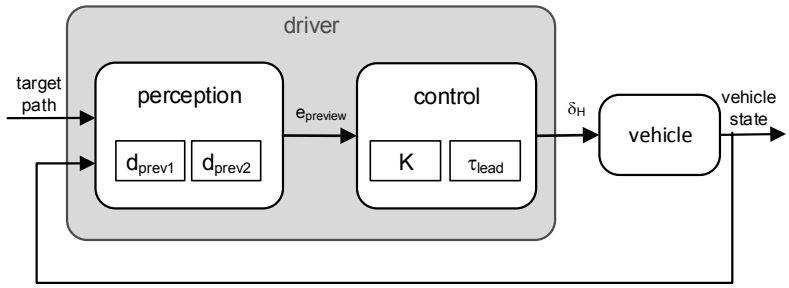


Figure 106. Driver model B with four adaptable driver parameters.

For driver model B no feedforward information is used.

Driver model C

Driver model C is equal to driver model B, but with additional feedforward information implemented. This feedforward information models the driver's experience. Before driving the measured double lane changes with different tires, drivers practiced several double lane changes to become accustomed to the changed vehicle handling behavior, with the objective to rule out learning effects during the measured double lane changes. Drivers explained that

practicing gave them information on how to perform the best steering control for these tires, *i.e.* when to start steering to the next lane, how fast to steer and with what amplitude. With this experience, the measured double lane changes are driven. This driver experience is incorporated in driver model C with a feedforward part that models the basic steering control necessary for performing a double lane change with two consecutive and opposite sine waves. Fig. 107 shows the implementation of driver model C.

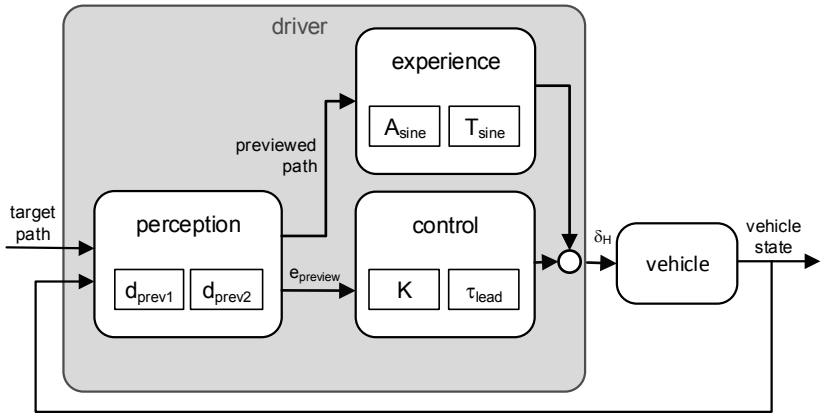


Figure 107. Driver model C with six adaptable driver parameters.

The sine waves start at the moment the driver focuses on the next lane target path, so d_{prev1} and d_{prev2} are input for this feedforward part. As this feedforward signals only models the basic steering control and to limit the number of additional driver parameters, both sine waves have the same parameters, period T_{sine} and amplitude A_{sine}. This feedforward steering wheel control signal is added to the feedback steering wheel control signal to form the total steering wheel control signal δ_H. Fig. 108 shows this feedforward steering wheel control signal of two sine waves, together with the feedback steering wheel control signal and the sum of these two signals, giving the total steering wheel control signal.

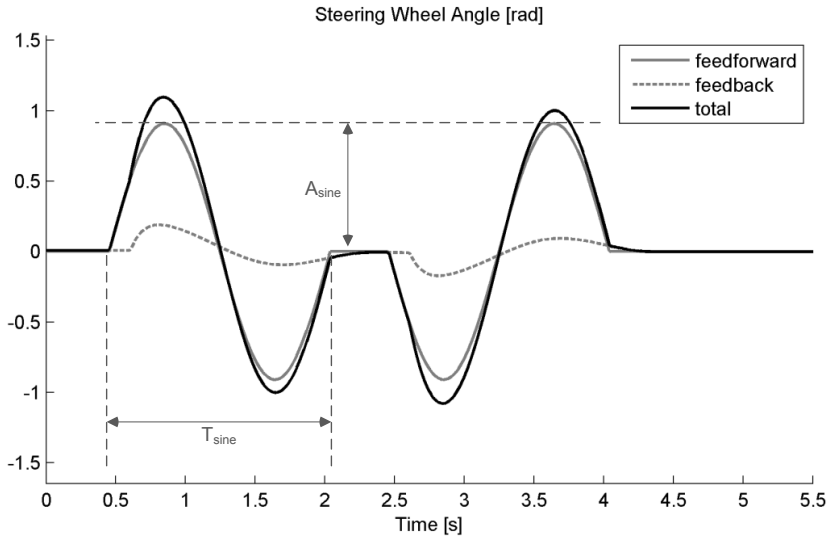


Figure 108. Driver model C steering wheel angle control signal consisting of feedforward and feedback part.

As this is a feedforward signal based on the driver's experience there is no influence of the driver's delay time τ. This influence is visible in the feedback part of the steering wheel control signal, although the preview error increases due to the change in previewed path, which is the time where the feedforward sine wave starts, the feedback steering wheel control will change only after the driver's delay time τ of 0.15 s.

Driver Models Parameter Estimation Method

For each double lane change, the adaptable parameters of each driver model A, B and C are estimated by taking the optimal values for these parameters so that the model output fits the measured output best in the least squares sense. Eq. (30) to (32) show the driver model parameter vectors, for respectively driver model A, B and C, that are estimated.

$$\boldsymbol{p_A} = \begin{bmatrix} t_{preview} \\ K \\ \tau_{lead} \end{bmatrix} \tag{30}$$

$$\boldsymbol{p_B} = \begin{bmatrix} t_{prev1} \\ t_{prev2} \\ K \\ \tau_{lead} \end{bmatrix} \tag{31}$$

$$\boldsymbol{p_C} = \begin{bmatrix} t_{prev1} \\ t_{prev2} \\ K \\ \tau_{lead} \\ T_{sine} \\ A_{sine} \end{bmatrix} \tag{32}$$

The model output is derived by simulating the driver model in closed-loop with the vehicle model with as model inputs the target path and the measured longitudinal velocity of the vehicle v_x. The model output signals that are fitted on the measured data are the steering wheel angle δ_H and the driven path of the vehicle Y. The steering wheel angle is the driver model output and the measured signal coming from the driver during testing. In addition, the driven path has been chosen as fit signal, because the model of the total driver-vehicle system should also represent real-world behavior. Because of the aforementioned dependency of the driver parameters, the set of optimized parameters is not unique. By fitting these two signals at the same time to the measured signals, the variation in the set of optimized driver parameters is reduced to values representing more realistic behavior.

The parameter optimization is done by calculating an error function F for every maneuver. The error function F, given in Eq. (33), is a vector valued function depending on the driver model parameter vector p, which calculates the difference between the model output and the test output for every time step for the steering wheel angle δ_H and the driven path Y. As both signals carry the same amount of information, they are scaled to the same level (Hyötyniemi, 2003), so these normalized errors of both signals are equally weighted. The errors are normalized by dividing them by the standard deviation of the test signal, calculated over all maneuvers, which also enables comparison of the error function between maneuvers. Time steps during 7 seconds are taken, starting when the vehicle's location is 8 meter before the first cone of the first lane. This time frame incorporates all the dynamics of driving the double lane change. The sampling frequency of the measured signals being 100 Hz, this resulted in 701 samples for each fitted output signal, resulting in 1402 samples for every double lane change. The error function F is defined as:

$$F(\boldsymbol{p}) = \begin{bmatrix} F_1(\boldsymbol{p}) \\ \vdots \\ F_{701}(\boldsymbol{p}) \\ F_{702}(\boldsymbol{p}) \\ \vdots \\ F_{1402}(\boldsymbol{p}) \end{bmatrix} = \begin{bmatrix} \left(\delta_{H,model,1}(\boldsymbol{p}) - \delta_{H,test,1}(\boldsymbol{p})\right) / \sigma_{\delta_H,test} \\ \vdots \\ \left(\delta_{H,model,701}(\boldsymbol{p}) - \delta_{H,test,701}(\boldsymbol{p})\right) / \sigma_{\delta_H,test} \\ \left(Y_{model,1}(\boldsymbol{p}) - Y_{test,1}(\boldsymbol{p})\right) / \sigma_{Y,test} \\ \vdots \\ \left(Y_{model,701}(\boldsymbol{p}) - Y_{test,701}(\boldsymbol{p})\right) / \sigma_{Y,test} \end{bmatrix} t \quad (33)$$

with driver model parameter vector p belonging to driver model A, B or C as given in respectively Eq. (30) to (32), $\delta_{H,model,n}$ the n^{th} sample of the driver model output, being the steering wheel angle value, $\delta_{H,test,n}$ the n^{th} sample of

the measured steering wheel angle value, $Y_{model,n}$ the n^{th} sample of the vehicle model Y position, $Y_{test,n}$ the n^{th} sample of the measured vehicle Y position, $\sigma_{\delta_H,test}$ the standard deviation of the measured steering wheel angle, signal, $\sigma_{Y,test}$ the standard deviation of the measured Y position of the vehicle.

The optimization of the driver model parameter vectors is done for every maneuver and for every driver model by minimizing the normalized sum squared error (SSE):

$$\min_{p} \|\boldsymbol{F}(\boldsymbol{p})\|_2^2 = \min_{\vec{p}} (F_1(\boldsymbol{p})^2 + \cdots + F_{1402}(\boldsymbol{p})^2) \tag{34}$$

This minimization is done with a nonlinear least squares optimization function lsqnonlin (MathWorks, 2014) using the Levenberg–Marquardt algorithm. To improve finding the global minimum, instead of a local minimum, this function is executed iteratively for every maneuver with several different starting points and upper and lower boundaries for the search space.

This driver model parameter estimation method resulted in so-called optimal model parameters for each driver model A, B and C for each maneuver.

8.2.3 Model Output Compared with Measured Output

The driver models are simulated with the optimized driver model parameters in closed-loop with the vehicle model with as model inputs the driver target path and the measured longitudinal vehicle velocity v_x. The driver model parameters are optimized by fitting the model output signals for the steering wheel angle δ_H and the driven path of the vehicle Y on the measured data. For this driver model method, it is important that the driver models produce realistic driver behavior, to let the chosen model parameters reflect the driver characteristics they stand for. To see if the driver models are able to produce realistic driver behavior, this model output is compared with the actual, measured output. Fig. 109 shows the mean values of the driver model A, B and C fitted output data and measured data for the maneuvers driven with the (extreme) combination of tire 6 driven by the professional drivers for the high speed demand. In addition, the normalized SSE value given in Eq. (34) is given for all three models. This shows that model C fits the combination of steering wheel angle and target path the best (in least squares sense), followed by driver model A and driver model B. All other combinations of driver group, tire and speed show similar behavior for the models.

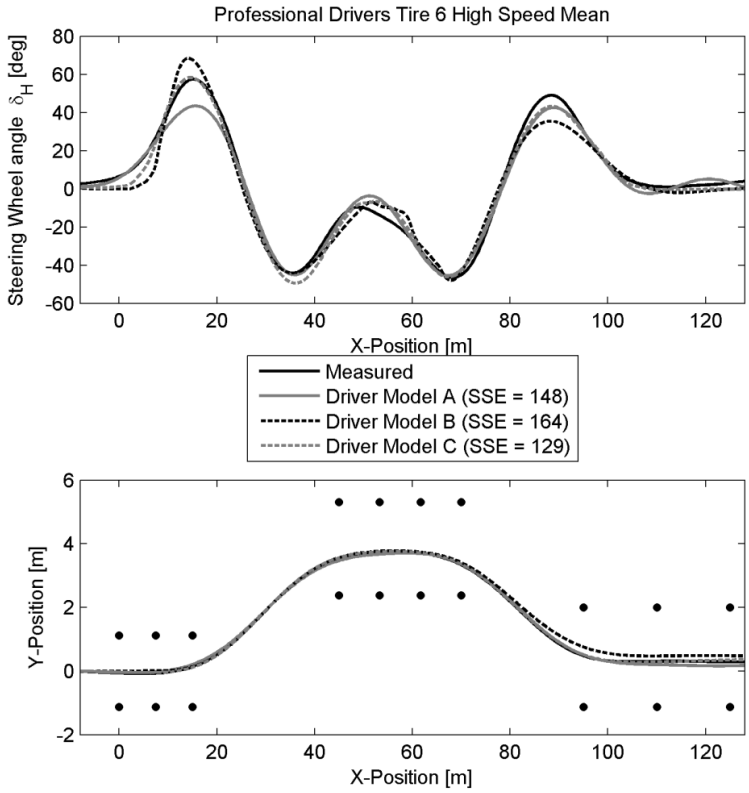


Figure 109. Mean values of the in closed-loop simulated driver models A, B and C output data and measured data for the maneuvers driven with tire 6 by the professional drivers for the high speed demand.

All the driver models with the optimized parameters follow the path closely, only driver model B is deviating from the measured path slightly in the last lane. The general behavior of the steering wheel angle is also followed, but deviations can be seen in the first and last peak values. Changes in driver parameter values to improve this, cause different behavior for the whole maneuver, which decreases the overall performance, measured in SSE. Apparently, the sacrifice to have a relative large error in the first and last peak value is still beneficial for the overall behavior, while the path is followed closely. For example, for driver model A to improve the behavior of the first and last peak, the driver parameter gain could be increased, the steering wheel angle will get larger for the same preview error. This behavior will be continued over the whole maneuver, giving increasing peak values in between and not following the path so accurate anymore.

8.3 Results and Discussion

In this section, the driver model parameters are presented related to the independent variables drivers group, speed and tire and the results are

discussed. This is done separately for each driver model A, B and C, followed by a summary for all three driver models. As is discussed in section 4.7.2 and validated in section 5.3, the variables speed and tire are manipulated to influence the driver's task demand and therefore the driver's mental workload, while this is done in the task related effort region A3. A driver model parameter is considered *speed dependent*, if the mean values for low speed are all higher or lower than the mean values for high speed. High speed corresponding with high driver's mental and physical workload and low speed with low mental and physical workload (as validated in section 5.3). A driver model parameter is considered *tire dependent*, if the mean values monotonically increase or decrease for respectively tire 2, 5 and 6, for both speed demands, corresponding to decreasing mental workload for the tires. If a driver parameter is found to be tire dependent in this experiment, the assumption that it depends on driver's mental workload, is validated.

8.3.1 Driver Model A

Fig. 110 to 112 show the results for the mean value of the three adaptable driver parameters, grouped by drivers group, tire and speed.

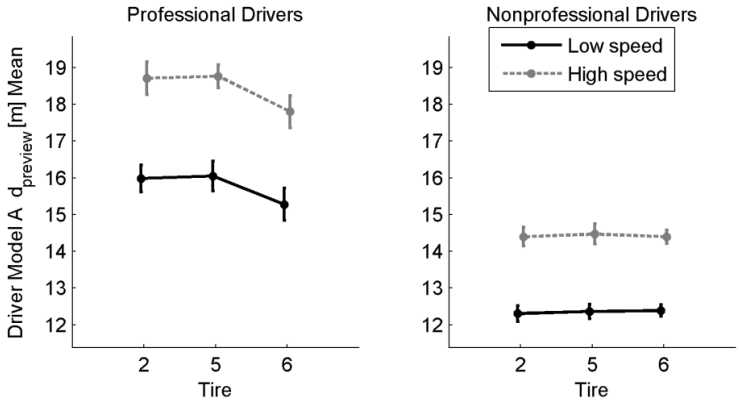


Figure 110. Mean values of the preview distance $d_{preview}$ of driver model A, grouped by drivers group, tire and speed.

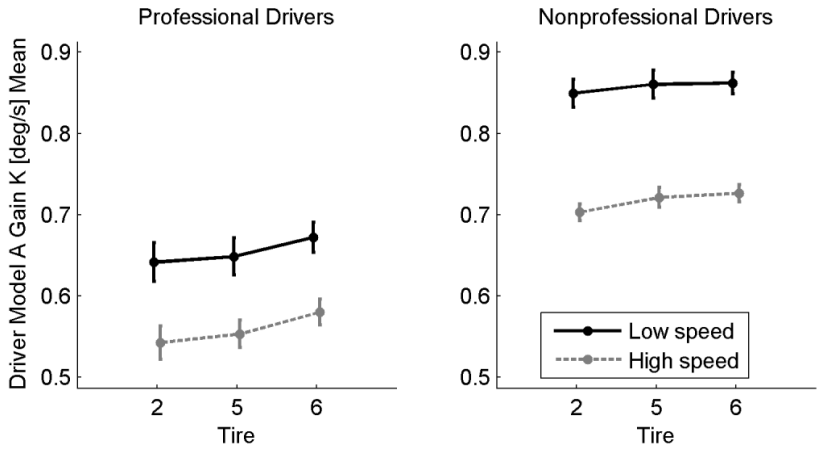


Figure 111. Mean values of the gain K of driver model A, grouped by drivers group, tire and speed.

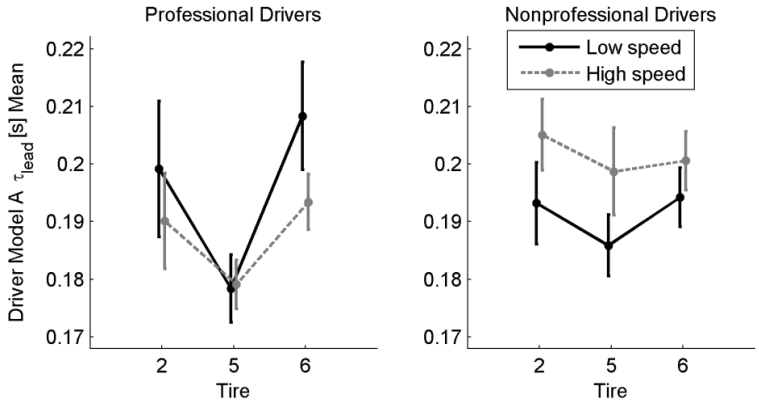


Figure 112. Mean values of the lead term time constant τ_{lead} of driver model A, grouped by drivers group, tire and speed.

The figures show that all the driver model A parameters are speed dependent. The preview distance values are higher for higher speed, which is consistent with previous research (section 8.1.3). The professional drivers have larger preview times for all conditions compared to the nonprofessional drivers, but the actual values for the low and high speed were also higher for the professional drivers (Table 5, page 62). Fig. 113 shows the results for the mean value of the driver preview time, grouped by drivers group, tire and speed.

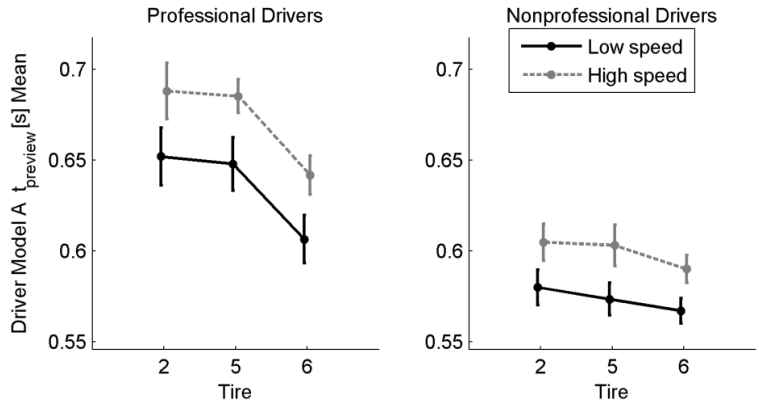


Figure 113. Mean values of the preview time $t_{preview}$ of driver model A, grouped by drivers group, tire and speed.

This reveals that the preview time is not constant, but also depending on the speed. The driver's look ahead distance, which for constant speed is equal to the speed times the preview time, increases, apparently compensating more than only the increasing speed. Speed dependency of the lead term time constant is less prominent, almost disappearing for the professional drivers.

Tire dependency of the parameters is weak, the values are close to each other and the SEM error bars have considerable overlap. An almost constant value can be seen for the preview distance and gain. The lead term time constant shows some tire dependency, but not monotonically, tire 5 having a minimum value. This lead term time constant was expected to be the most related to workload based on literature and previous research (see section 2.4.1, 2.4.2 and 8.1.3), but this is not validated from these results of this driver model.

Looking at the combination of parameters, overall, a higher preview time is combined with a lower driver gain, consistent with their assumed mutual dependency (as explained in section 8.2.2 and (Pauwelussen, 2012)). The lead term time constant shows a lower value for tire 5, compared to tires 2 and 6, which have comparable values, especially for the professional drivers. This relative low value of τ_{lead} for tire 5 is accompanied with a relative high value (compared to a linear trend) for $t_{preview}$. Also in this case, mutual parameter dependency could explain this: an slight increased look ahead distance could imply the driver has to anticipate less.

8.3.2 Driver Model B

Fig. 114 to 117 show the results for the mean value of the four adaptable driver parameters, grouped by drivers group, tire and speed.

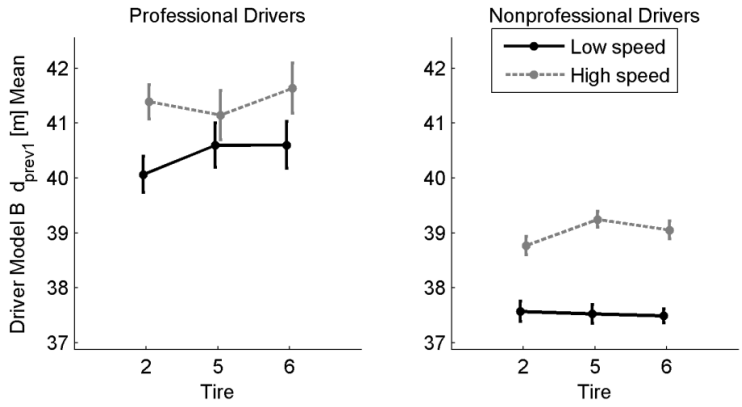


Figure 114. Mean values of the preview distance d_{prev1} of driver model B, grouped by drivers group, tire and speed.

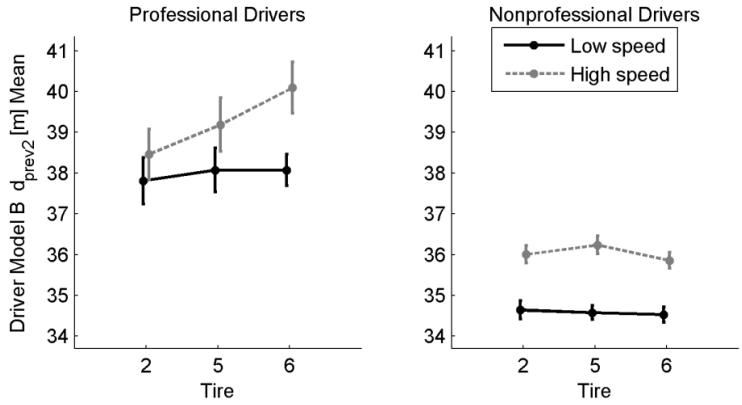


Figure 115. Mean values of the preview distance d_{prev2} of driver model B, grouped by drivers group, tire and speed.

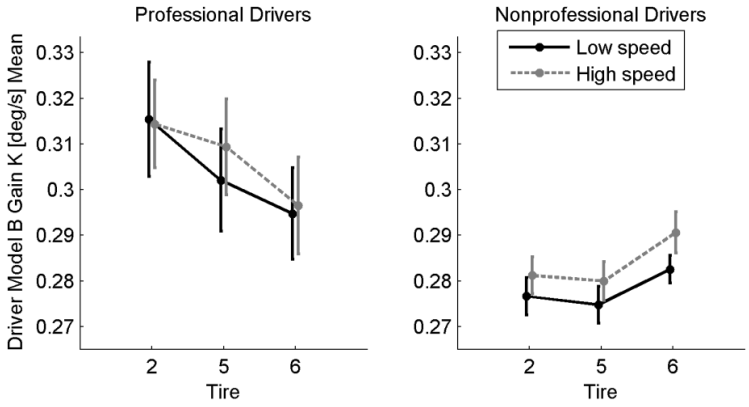


Figure 116. Mean values of the gain K of driver model B, grouped by drivers group, tire and speed.

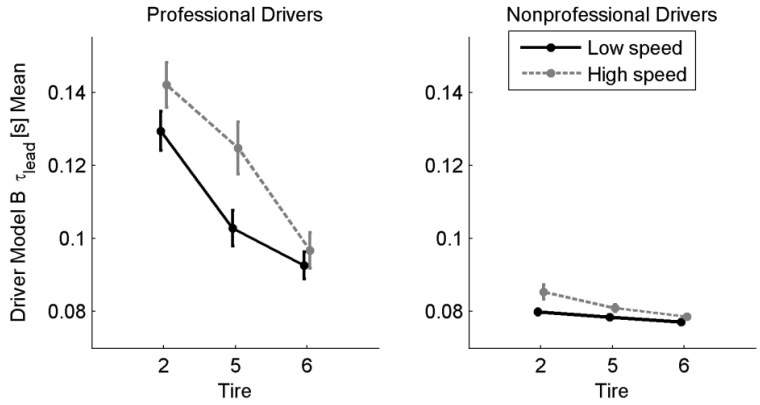


Figure 117. Mean values of the lead term time constant τ_{lead} of driver model B, grouped by drivers group, tire and speed.

Also for this driver model B the figures show speed dependence for almost all model parameters. Also for this driver model B a higher speed gives a larger preview distance. The values found for the preview distances are more than twice the values found for driver model A, but these values do not represent similar concepts. For driver model A this is a constant distance which the driver looks ahead on a (imaginary) continuous path he should follow during the whole maneuver, while for driver model B these are discrete points on the current lane where the driver switches from looking at the current lane to looking at the next lane.

The increase in preview distance is not compensating fully the increase in speed. This can be seen from the preview times, which are lower for higher speed, in contrast with driver model A, as can be seen in Fig. 118 and 119.

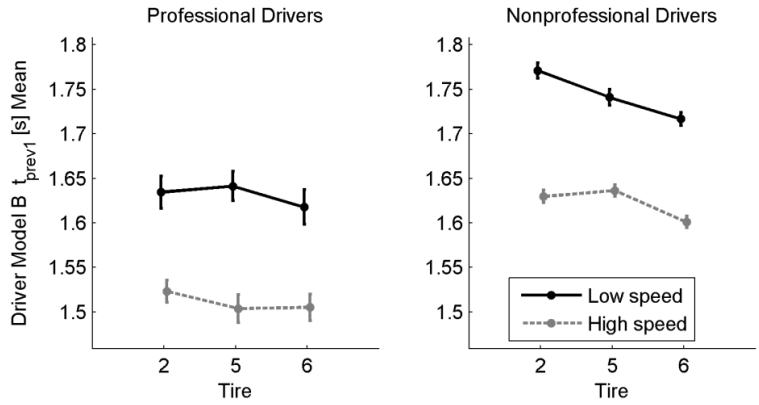


Figure 118. Mean values of the preview time t_{prev1} of driver model B, grouped by drivers group, tire and speed.

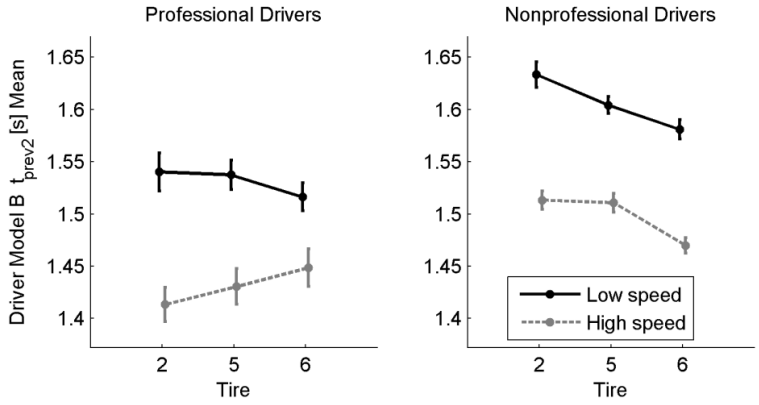


Figure 119. Mean values of the preview time t_{prev2} of driver model B, grouped by drivers group, tire and speed.

An explanation is that the preview distance already is long at low speeds and further increase of the distance does not improve the performance. Fig. 116 shows that the gain is almost not decreased for high speed and corresponding larger preview distance, but kept the same (professional drivers) or is increased (nonprofessional drivers). This suggests that the driver, not able to compensate fully for the higher speed with increasing preview distance, compensates by not decreasing, or even increasing his gain.

The parameters have differences in mean values between tires, but no tire dependency consistent over the speed is shown. Tire dependency can be seen for the driver gain parameter for the professional drivers, the better the handling, the lower the gain, although the error bars have quite an overlap. This could be seen as the better handling tire requiring less steering wheel input. However, this effect is not seen for the nonprofessional drivers, having an increasing in gain for tire 6. Tire dependency of the lead time constant is strong, especially for the professional drivers. A tire dependency can be seen and the error bars are small and show no overlap. The better the handling tire, the lower the lead time. Interpreting lead term as anticipation, means a better handling tire requires less driver anticipation. For the nonprofessional drivers this effect is not so prominent. This could be caused by the simulation model having a lower boundary for the lead term time constant , *i.e.* the value of the lag term time constant. This is done, because a lead term time constant value lower than the lag term time constant value would not result in a lead term effect (anticipating), because it is compensated by the lag term. The lag time constant was set to a constant value of 0.08 s as explained in section 8.2.2. This implied that the estimation algorithm could not get a lower total error function if the lead term time constant was lower than 0.08 s. Although for this nonprofessional drivers group the tire dependence of the low speed demand could not be validated, the high speed demand does show the same trend: lower values for better handling tires.

8.3.3 Driver Model C

Fig. 120 to 125 show the results for the mean value of the six adaptable driver parameters, grouped by drivers group, tire and speed.

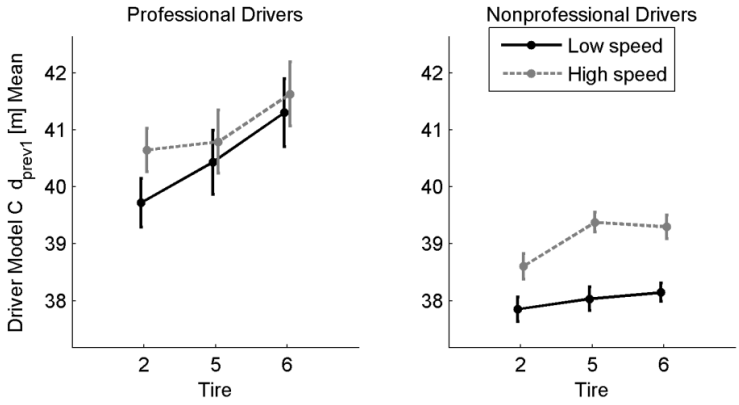


Figure 120. Mean values of the preview distance d_{prev1} of driver model C, grouped by drivers group, tire and speed.

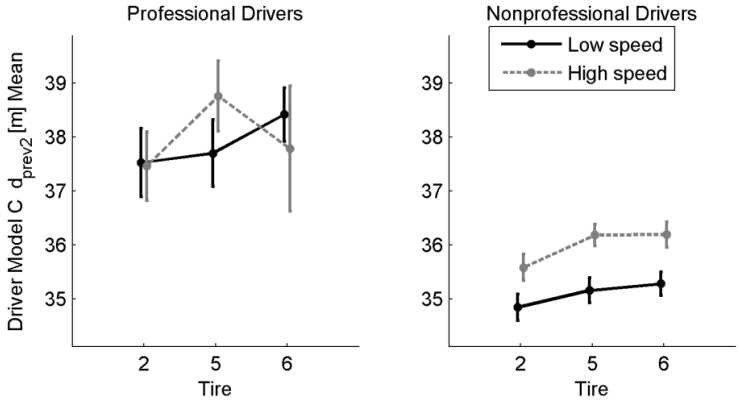


Figure 121. Mean values of the preview distance d_{prev2} of driver model C, grouped by drivers group, tire and speed.

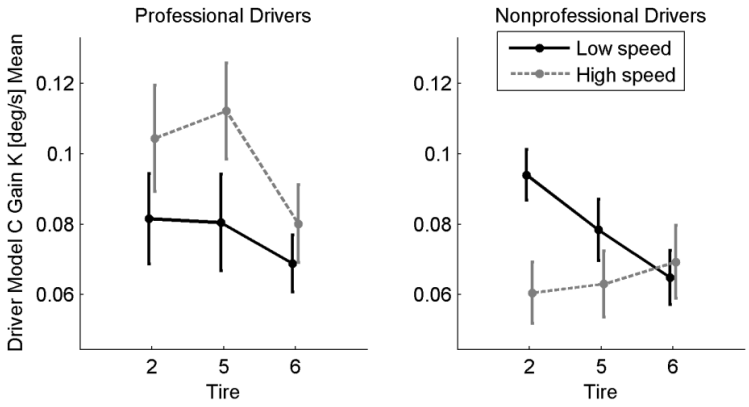


Figure 122. Mean values of the gain K of driver model C, grouped by drivers group, tire and speed.

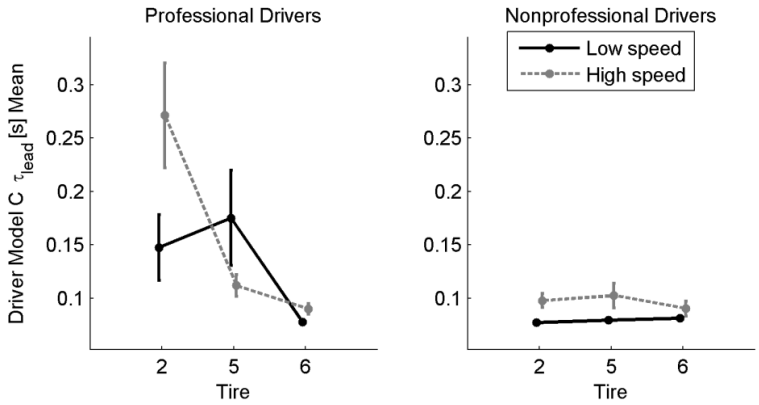


Figure 123. Mean values of the lead term time constant τ_{lead} of driver model C, grouped by drivers group, tire and speed.

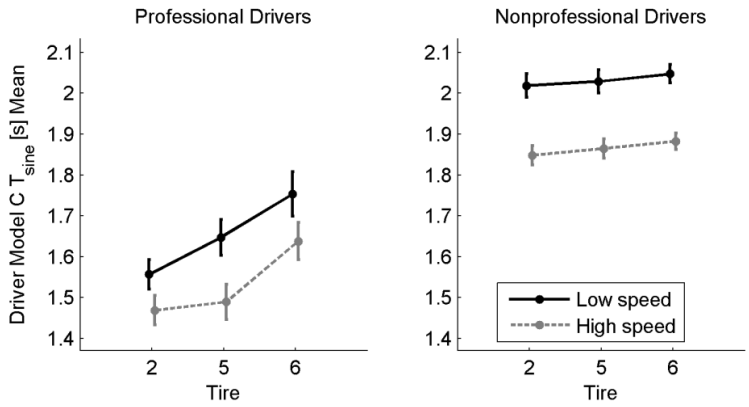


Figure 124. Mean values of the period T_{sine} of driver model C, grouped by drivers group, tire and speed.

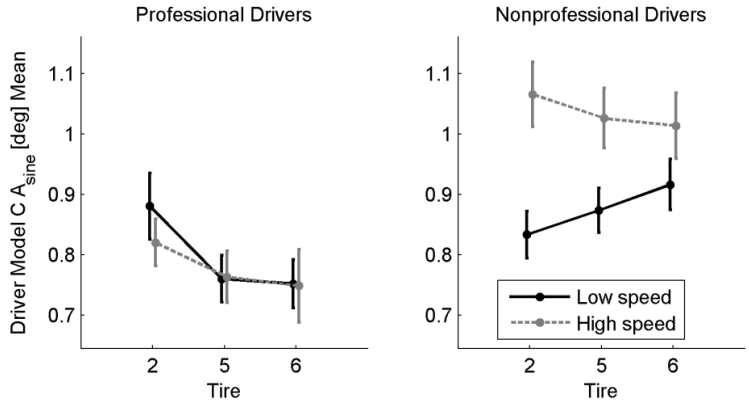


Figure 125. Mean values of the amplitude A_{sine} of driver model C, grouped by drivers group, tire and speed.

The results for driver model C show varying speed dependence of the model parameters. The preview distances show similar values as for driver model B, although the values are less speed dependent for the professional drivers. The preview times are speed dependent for both driver groups, with similar values as driver model B as can be seen in Fig. 126 and 127.

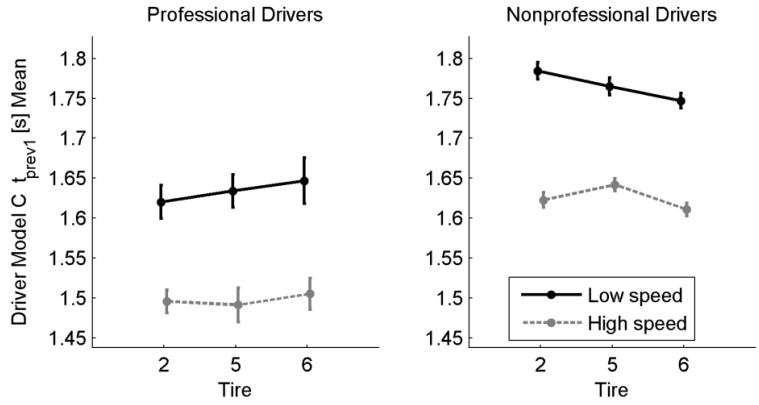


Figure 126. Mean values of the preview time t_{prev1} of driver model C, grouped by drivers group, tire and speed.

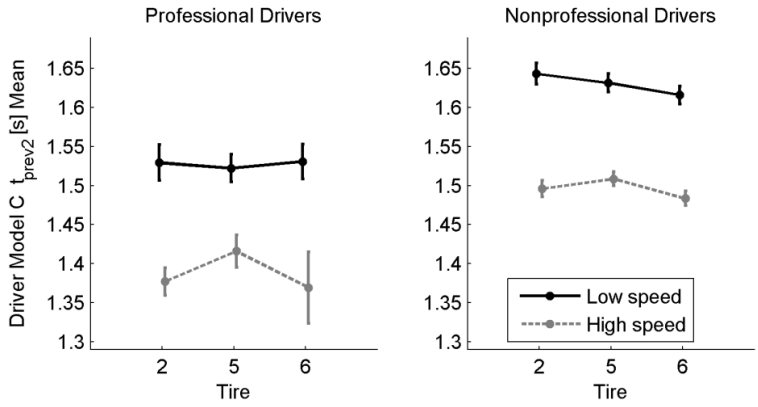


Figure 127. Mean values of the preview time t_{prev2} of driver model C, grouped by drivers group, tire and speed.

The gain values are all much lower than for driver model B, as could be expected, because there is a feedforward part in the steering wheel angle in model C that forms a large part of the necessary steering wheel angle to follow the path and which is not depending on the path error. Therefore, the driver has to perform smaller additional steering to correct path errors, resulting in this lower driver gain. The speed dependence of the gain is different for the driver groups, the professional drivers have higher gain for higher speed, the nonprofessional drivers have higher gain for lower speed, except for tire 6. The lead term time constant shows different mean values for the both speeds, but not consistent higher or lower for the professional drivers. The feedforward driver parameter period T_{sine} is lower for higher speed, as is expected; driving with a higher speed on the same maneuver requires faster steering. Speed dependence of the feedforward amplitude A_{sine} is only seen for the nonprofessional drivers, being higher for high speed. This parameter is assumed to be related to the gain, both parameters determine how much the driver steers. The results show that a higher value for A_{sine} for the nonprofessional drivers is combined with a lower value for the gain. In contrast, the professional drivers have a higher gain, but almost equal values for A_{sine}. This suggests these two driver parameters can compensate each other.

Some tire dependency can be seen for the preview distances, but not consistent over the speed and groups. This can also be concluded for the gain parameter. Also here the mutual dependency of this parameter with A_{sine} can be seen, a relative high value of the gain is combined with a relative low value of A_{sine}. The lead term time constant does not show the same tire dependency as with driver model B. The mean values for the professional drivers for low speed show no monotonically in- or decrease and the error bars for the professional drivers are larger. The period T_{sine} shows tire dependency for the professional drivers, but almost constant values for the nonprofessional drivers.

8.3.4 Summary

The results and discussion for the three driver model parameters are summarized here. For every driver model parameter the speed dependency is given by (Table 14.)

- An evaluation symbol, representing the mean driver model parameter values for the tires for low speed with respect to high speed.
- A background color, representing error bars the driver model parameters for a certain speed.

Table 14. Evaluation symbols and ackground color for the speed dependency of the driver model parameters.

Symbol	Mean Values for Low Speed w.r.t. High Speed
+	All higher or lower
–	Not all higher or lower

Background color	Error Bars for a Certain Speed
	No error bars overlap
	One error bar overlaps
	More error bars overlap
	Not relevant

Table 15 gives an overview of the parameters speed dependency evaluated for both drivers groups separately and for both groups together. In addition, the preview time parameters are included, because their speed dependency can differ from the preview distance speed dependency.

Table 15. Summary of Speed Dependency of the Driver Model Parameters grouped by driver group.

Driver Model	Parameter	Speed Dependency		
		Professional driver group	Nonprofessional driver group	All drivers
A	preview distance $d_{preview}$	+	+	+
	preview time $t_{preview}$	+	+	+
	gain	+	+	+
	lead time constant τ_{lead}	–	+	–
B	preview distance 1 d_{prev1}	+	+	+
	preview distance 2 d_{prev2}	+	+	+
	preview time 1 t_{prev1}	+	+	+
	preview time 2 t_{prev2}	+	+	+
	gain K	–	+	–
	lead time constant τ_{lead}	+	+	+
C	preview distance 1 d_{prev1}	+	+	+
	preview distance 2 d_{prev2}	–	+	–
	preview time 1 t_{prev1}	+	+	+
	preview time 2 t_{prev2}	+	+	+
	gain K	+	–	–
	lead time constant τ_{lead}	–	+	–
	period T_{sine}	+	+	+
	amplitude A_{sine}	–	+	–

This table shows that preview time is a strong speed dependent parameter for all driver models, having all values higher or lower for different speed and no overlap in error bars. This is also seen for the gain and preview distance of driver model A and the period T_{sine} of driver model C. The preview distances for the other models show strong speed dependence only for the nonprofessional drivers. The lead term time constant shows only speed dependency for driver model B, being strong for the nonprofessional drivers.

For every driver model parameter the tire dependency is evaluated in the same way as the speed dependency. Table 16 presents the meaning of the evaluation symbols for the tire dependency, the background color meaning of Table 14 is also used here.

Table 16. Evaluation Symbols for the Tire Dependency of the Driver Model Parameters.

Symbol	Mean Values for Tires 2, 5 and 6 for Both Speeds
+	Increasing or decreasing
–	Not increasing or decreasing

Table 17 gives an overview of the parameters tire dependency evaluated for both drivers groups separately and for both groups together, including the preview time parameters.

Table 17. Summary of Tire Dependency of the Driver Model Parameters grouped by driver group.

Driver Model	Parameter	Tire Dependency		
		Professional driver group	Nonprofessional driver group	All drivers
A	preview distance $d_{preview}$	–	–	–
	preview time $t_{preview}$	+	+	+
	gain	+	–	–
	lead time constant τ_{lead}	–	–	–
B	preview distance 1 d_{prev1}	–	–	–
	preview distance 2 d_{prev2}	–	–	–
	preview time 1 t_{prev1}	–	–	–
	preview time 2 t_{prev2}	–	–	–
	gain K	+	–	–
	lead time constant τ_{lead}	+	(+)	+
C	preview distance 1 d_{prev1}	+	–	–
	preview distance 2 d_{prev2}	–	+	–
	preview time 1 t_{prev1}	+	–	–
	preview time 2 t_{prev2}	–	–	–
	gain K	–	–	–
	lead time constant τ_{lead}	–	–	–
	period T_{sine}	+	+	+
	amplitude A_{sine}	+	–	–

Comparing this table to the table of the speed dependencies, it can be seen that the overall evaluation of the tire dependencies is less good than the speed dependencies. This was to be expected based on the setup of the experiment. The different speed demands were applied to give large differences in driver's task demand and therefore large differences in driver's mental workload, the differences between tires being smaller. Validation of this assumption (section 5.3) confirmed this for the nonprofessional drivers, but for the professional drivers the differences were similar. This is confirmed by this table: overall tire dependencies of the parameters are seen more for the professional drivers than for the nonprofessional drivers. Where the preview times all showed strong speed dependency, this is not seen for tire dependency, only for driver model A the preview time shows some tire dependency. The lead term time constant of driver model B is having tire dependency for all drivers. The value for the nonprofessional drivers being put between parenthesis, because this results from the evaluation, but probably due to reaching the lower bound of this parameter, the error bars are almost zero. The period T_{sine} shows tire dependency, being again stronger for the professional drivers.

8.4 Conclusions

How a driver perceives the tire handling behavior was the main question to be answered in this chapter. The driver model method answered this question by relating tire handling assessment scores to driver's mental and physical workload and analyzing the optimized driver model parameters representing

driver characteristics like preview time, gain, lead and lag behavior and from this the adapting behavior of the driver. Because of the dependency of the driver parameters, the set of optimized parameters is not unique. By fitting the model steering wheel angle and the driven path of the vehicle at the same time to the measured signals, the variation in the set of optimized driver parameters is reduced to values representing realistic behavior.

To capture complex driver behavior in a few adaptable driver model parameters is no easy task. Driver models A, B and C, used in this driver model method, have an increasing number of adaptable driver model parameters. The more adaptable parameters, the better the model is able to represent the actual driver steering behavior. Analyzing the optimized model parameters revealed that all models could distinguish well between low and high mental and physical workload, seen for the low and high speed demand. Many of the parameters showed clear speed dependence. This was less clear for the workload differences due to different tires. Tire dependency was seen more for the professional drivers than for the nonprofessional drivers, which is in line with the fact that the professional drivers showed larger differences in workload between tires.

The lead term constant for driver model B showed to be the best candidate for both drivers groups, having good tire dependency. Although mutual dependent driver model parameters represent actual driver behavior and are present in all three driver models, too many drive model parameters can make the results less clear. Model A, using a continuous preview path, and model C, having the most adaptable driver model parameters, both seemed to have too much mutual dependencies in their parameters. Driver model B seems to be a good compromise, having a simple, realistic perception part using a discrete preview path, but also having less driver model parameters than model C.

Analysis of the model parameters showed how the driver adapts his perception and control to different workload tasks induced by different speeds and different tires. This also revealed the dependencies between model parameters, which gave insight in dependencies in driver behavior. In addition, differences between the professional drivers and nonprofessional drivers could be explained by looking at the corresponding driver parameters.

This driver model method shows how different driver behavior due to different tires can be explained and how this can be related to the driver’s experienced mental and physical workload.

9. Overall Conclusions and Discussion

9.1 Overall Conclusions

This research was done to gain more knowledge about subjective tire handling assessment by professional test drivers with the aim of improving tire handling assessment methods. This supports the development of good handling tires, contributing to the active safety of a vehicle and to the pleasure of driving. This main goal is detailed into the following research objectives:

- To provide information on *what* handling behavior is considered as good by the driver, by predicting subjective tire handling assessments based on derived objective measures and by analyzing the these measures. This information can support the actual driver's subjective evaluations and provide comparative information for assessments.
- To provide information on *how* the driver's subjective assessment is formed by analyzing driver model parameters from vehicle handling simulations.

Professional tire test drivers are very capable of doing their work; improvement is therefore searched for in the assessment methods. Specifically, by focusing on the driver, because handling is determined by the closed-loop driver-vehicle system and the driver is the core of subjective assessment.

Three methods, all based on field experiments, are derived which fulfil the research objectives: using a GRNN, based on workload measures and based on driver models. They have in common that they predict the driver's subjective assessment of tire handling, based on vehicle dynamics measurements. The differences lie in the way they derive and utilize these measurements.

The GRNN method (described in chapter 6) and workload measures method (described in chapter 7) provide information on the 'what'-question:

- The GRNN showed good prediction of the driver's subjective assessment scores covering driving a handling track, based on vehicle dynamic measurements taken only during the double lane changes.
- The GRNN can even perform well on predicting handling aspects not related to driving the double lane change.
- Analysis of the important measures used by the GRNN can provide information on the vehicle dynamics behavior relevant for the driver. For general handling behavior, regarding all aspects, this showed that metrics based on lateral velocity, steering wheel moment and vehicle slip angle were most relevant.

- Additional advantages of using a GRNN compared to other regression methods are: it works well with a limited dataset, does not need a-prior knowledge about the data or underlying model, has good prediction between data points, behaves well for extrapolation. In addition, constructing the GRNN is easy and adding new examples, with possibly different in- and outputs can be done during usage, incorporating new knowledge when it becomes available.
- Mental workload measures are found to be indicators of driver's perceived tire handling behavior, even when performance measures do not show differences. More mental workload indicates less tire handling performance, the driver has to put in more effort to keep the performance high.
- HFA showed to be a promising measure in this research for indicating driver's mental workload during tire handling, but needs more research to make these results broader applicable.
- Steering wheel measure HFA showed to be the best mental workload measure for this research.
- Tire handling assessment is found to be situated in region A3 of the performance - workload model of De Waard (Waard, de, 1996) (Fig. 19), which let driver's workload measures distinguish between tire handling differences when performance measures cannot.
- The relationship between performance and driver's workload in region A3 of the performance - workload model of De Waard (Waard, de, 1996) is found to be different for tire handling (Fig. 47). For tire handling, increasing task demand gives linearly increasing mental workload, but at a certain point the mental workload increases rapidly to maximum workload, just when the performance decreases.

The driver model method (described in chapter 8) provides information on the 'how'-question:

- The parameters of the driver models applied in this research represent driver characteristics. By analyzing the identified driver model parameters, information on these related driver characteristics was derived. Especially the parameter changes due to different task demands and corresponding mental and physical workload provides information on the adapting behavior of the driver.
- From the three derived driver models A, B and C, driver model B showed to be the best model by having a simple, realistic perception part using a discrete preview path, but also having the optimal number of driver model parameters to explain driver behavior.
- The lead term constant of driver model B showed to be the best candidate for predicting subjective tire handling for both professional and nonprofessional drivers.

9.2 Discussion

This research is a first step in predicting and quantifying the driver's tire feeling. This must provide more knowledge about subjective tire handling assessment by professional test drivers with the aim of improving tire handling assessment methods. With a growing demand for tires, a tire manufacturer is faced with a growing need for tire subjective assessments, whilst limited resources are available. The methods presented here offer new possibilities to gather this information on *what* handling behavior is considered as good by the driver, and *how* the driver's subjective assessment is formed.

From a scientific point of view, the "how" question is probably the most interesting. It gives a possible explanation for the driver's perception and adaptation during tire handling assessment. Although the first driver research started in the late fifties of the previous century, and handling was soon incorporated, driver research related to tire handling is very limited and often focusing on the influence of tire characteristics on the vehicle-part of handling.

A possible reason could be, that for tire testing, performance based measures can show equal results, while the driver subjective results are different. This research found that tire handling assessment is situated in the task-related effort region of the performance - workload model of De Waard (Waard, de, 1996) (Fig. 19), where performance measures cannot distinguish between different task demand, like driving with different handling tires. Hence, differences can only be revealed when the focus is on the driver.

In this research, the driver-part of handling is therefore focused on. By using parameters representing relevant, realistic driver behavior, the driver model method is able to reveal how the driver's perception and adaptation during tire handling manifests itself. Relating these results to the tire handling assessment scores can also provide us with information on how the driver experiences good handling.

From a tire manufacturer point of view, the "what" question is probably the most interesting. It provides information that can support the actual driver's subjective evaluations and that can provide comparative information for assessments. The method using a GRNN can predict subjective tire handling assessments based on derived objective measures. Analyzing these measures can also provide what handling behavior is considered as good by the driver.

The value of the methods derived here for a tire manufacturer depends not only on promising results, but also on the ease of implementation. Nowadays, subjective assessment of tire handling is not only done by high speed driving on handling tracks like the Nürburgring, but more and more on automotive proving grounds like in Papenburg, including repeating standardized maneuvers, like lane changes, on different speeds. This makes the methods derived in this research, which are based on such maneuvers, easier to implement in everyday tire testing.

This research was based on field experiments derived from tire testing practice, to ensure validity of the results, but the number and complexity of the measurements used for vehicle dynamics and drivers was nowhere near representing real testing conditions where no or limited measurements were

taken. Fortunately, nowadays vehicle dynamics and driver measurement systems are getting better, more affordable and therefore used more often during handling assessments.

As mentioned in the chapters describing the different methods, based on the results, not all measurements taken for this research are needed for implementation. For optimal performance of the GRNN only a few of all the used measurements are needed. For the mental workload measures, the steering measure HFA was the most promising one, requiring only a steering wheel angle measurement. The driver model method could be implemented with measurements representing the signals that should be fitted. For this research, these signals were the steering wheel angle and driven path, but other signals could also be used, like lateral acceleration or yaw rate.

Robustness of the methods in the sense of changing tire testing methods in time, was specifically addressed in the chapter describing the GRNN. The GRNN can incorporate additional measurements, evaluation aspects and results from new assessments and these can also be removed. Although for the methods, most often the double lane change is taken as handling maneuver in the field experiments, the methods can be adapted to other maneuvers, because they work with differences between measures instead of absolute values. If changing the maneuver is necessary, is something that has to be well considered; the GRNN analysis showed that measures derived from the double lane change were even able to predict aspects related to other parts of the handling track.

When used for virtual testing, the results do not depend on the measurement availability, as in simulation all signals are available, but on the availability of representative tire and vehicle models. If these models are available, the methods can be applied as with real testing and could give a prediction of the driver's subjective assessment and mental and physical workload.

The internal reliability, the consistence of the results between the methods, is supported by the similar results given by the different methods for the same experiments. For example, for driving double lane changes, the results from the three different methods were in line with each other. The external reliability, the consistence of the results for new tests, is something that has to be researched further, before the methods can be implemented.

Additional research must also be done for application of heart rate measures, the results presented here are promising, but only based on one or two drivers.

Overall, this research provides a first step in opening up the 'black box' of the driver's assessment by quantifying the driver's tire feeling. However, knowing that this box will never be opened totally, we ourselves will always have the pleasure to 'feel the tire'.

References

Abe, M. (2009). *Vehicle handling dynamics: theory and application*. Butterworth Heinemann.

Abe, M., Kano, Y., & Shibahata, Y. (2009). Driver model based vehicle handling quality evaluation. *IAVSD*. Stockholm.

Apollo Vredestein BV. (2014). Designed to protect you. Retrieved December 1, 2014, from http://www.apollovredestein.nl/about-apollo-vredestein/corporate-social-responsibility?Lang=en-US

Arts, L. (2007). *Tyre test driver models. Literature study, model choice and model implementation to model tyre test drivers*.

Aurell, J., Fröjd, N., & Nordmark, S. (2000). Correlation between objective handling characteristics and subjective perception of handling qualities of heavy vehicles. *AVEC*.

Bakker, E., Nyborg, L., & Pacejka, H. B. (1987). Tyre modelling for use in vehicle dynamics studies. Society of Automotive Engineers,Warrendale, PA.

Bakker, E., Pacejka, H. B., & Lidner, L. (1989). A New Tire Model with an Application in Vehicle Dynamics Studies. doi:10.4271/890087

Bergman, W. (1973). Measurement and subjective evaluation of vehicle handling. *SAE Paper*, *730492*.

Besselink, I. J. M., Pacejka, H. B., Schmeitz, A. J. C., & Jansen, S. T. H. (2005). The MF-Swift tyre model: extending the Magic Formula with rigid ring dynamics and an enveloping model. *JSAE Review*, *26*(2).

Besselink, I. J. M., Schmeitz, A. J. C., & Pacejka, H. B. (2010). An improved Magic Formula/Swift tyre model that can handle inflation pressure changes. *Vehicle System Dynamics*, *48*(S1), 337–352.

BMW AG. (2014). Freude am Fahren. Retrieved December 1, 2014, from https://www.bmw.de/de/home.html

Cavanagh, P. R., & Komi, P. V. (1979). Electromechanical delay in human skeletal muscle under concentric and eccentric contractions. *European Journal of Applied Physiology and Occupational Physiology*, *42*(3), 159–163. doi:10.1007/BF00431022

Chen, D. C., & Crolla, D. A. (1998). Subjective and objective measures of vehicle handling: drivers & experiments. *Vehicle System Dynamics*, *29*(sup1), 576–597. doi:10.1080/00423119808969588

Chen, D. C., Crolla, D. A., Alstead, C. J., & Whitehead, J. P. (1996). A Comprehensive Study Of Subjective And Objective Vehicle Handling Behaviour. *Vehicle System Dynamics*, *25*(sup1), 66–86. doi:10.1080/00423119608969188

Cnossen, F. (2000). Adaptive strategies and goal management in car drivers. *Unpublished Doctoral Dissertation, University of Gronigen, Gronigen, The Netherlands*.

Cole, D. J., Pick, A. J., & Odhams, A. M. C. (2006). Predictive and linear quadratic methods for potential application to modelling driver steering control. *Vehicle System Dynamics*, *44*(3), 259–284.

Crolla, D. A., Chen, D. C., Whitehead, J. P., & Alstead, C. J. (1998). Vehicle handling assessment using a combined subjective-objective approach. *SAE Transactions*, *107*(6), 386–395.

Dixon, J. C. (1996). *Tires, Suspension and Handling, 2nd edition*. (J. C. Dixon, Ed.). SAE.

Donges, E. (1978). A two-level model of driver steering behavior. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *20*(6), 691–707.

Droogendijk. (2010). A new neuromuscular driver model for steering system development.

Dugoff, H., Fancher, P. S., & Segel, L. (1970). An analysis of tire traction properties and their influence on vehicle dynamic performance. *SAE Paper*, *700377*.

Edelmann, J., Plöchl, M., & Lugner, P. (2007). A driver model for vehicle system dynamics simulation. *Proceedings of the European Automotive Congress "Automobile for the Future"*.

Fiala, E. (1954). Seitenkrafte am rollenden Luftreifen. *VDI-Zeitschrift*, *96*(973), 26–41.

Field, A. (2009). *Discovering statistics using SPSS*. Sage Publications Limited.

Freedonia. (2012). *World Tires; Industry Study with Forecasts for 2015 & 2020*.

Genta, G. (1997). *Motor Vehicle Dynamics: Modeling and Simulation*. (G. Genta, Ed.)*Series on Advances in Mathematics for Applied Sciences vol. 43* (2003rd ed., Vol. 43). World Scientific Publishing Company.

Genta, G., & Morello, L. (2009). *The automotive chassis: Components design. Vol. 1* (Vol. 1). Springer.

Guger, C., Edlinger, G., Leeb, R., Pfurtscheller, G., Antley, A., Garau, M., … Slater, M. (2004). Heart-rate variability and event-related ecg in virtual environments (pp. 240–245). Citeseer.

Hacker, W. (2006). ILO Encyclopaedia of Occupational Health and Safety, Ch. 29, article: Mental Workload. (I. L. O. 2006, Ed.). Geneva.

Heissing, B., & Ersoy, M. (2011). *Chassis Handbook: Fundamentals, Driving Dynamics, Components, Mechatronics, Perspectives*. Vieweg+Teubner.

Hendrikx, J. (1997). *Numerical optimization of tire characteristics with respect to car handling*. Technische Universiteit Eindhoven.

Hill, R. (1987). Correlation of Subjective Evaluation and Objective Measurement of Vehicle Handling.

Hogema, J. H., & Veltman, J. A. (2002). Werkbelasting en rijgedrag tijdens duisternis: Eerste veldexperiment. *TNO-Rapport TM-02-C046. Soesterberg: TNO Technische Menskunde*.

Horiuchi, S., & Sunada, K. (1998). Synthesis of Driver Assistance System for Lane-Following Using Generalized Predictive Control. *AVEC*.

Horiuchi, S., & Yuhara, N. (1996). An Analytical Approach to the Prediction of Vehicle Handling Qualities Using Multi-Input Driver Model. *AVEC*.

Hyötyniemi, H. (2003). *Multivariate regression: Techniques and tools*. Helsinki University of Technologie.

ISO. (1991). IS0 8855: 1991 (E/F) Road vehicles - Vehicle dynamics and road-holding ability - Vocabulary.

ISO. (1999). ISO 3888-1:1999(E) passenger cars - Test track for a severe lane-change manoeuvre- Part 1: Double lane change.

ISO. (2003). ISO 7401:2003(E) Road vehicles - Lateral transient response test methods - Open-loop test methods.

ISO. (2004). ISO 4138:2004(E) Passenger cars — Steady-state circular driving behaviour — Open-loop test methods.

ISO. (2006). ISO 15037-1:2006(E) Road vehicles — Vehicle dynamics test methods — Part 1: General conditions for passenger cars.

Jagacinski, R. J., & Flach, J. (2003). *Control theory for humans: Quantitative approaches to modeling performance*. CRC.

Jahn, G., Oehme, A., Krems, J., & Gelau, C. (2005). Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. *Transportation Research Part F: ...*.

Jaksch, F. (1977). Driver-vehicle interaction with respect to steering controllability. *Vehicle System Dynamics*, *6*(2-3), 185.

Johansson, R. (2006). *System Modeling & Identification*. (P. Hall, Ed.).

Jürgensohn. (2007). Control Theory Models of the Driver.

Kamalakannan, B., Groves, W., & Freivalds, A. (2007). Predictive models for estimating metabolic workload based on heart rate and physical characteristics. *J. SH&E Res*.

Kasabov, N. K. (1996). *Foundations of neural networks, fuzzy systems, and knowledge engineering*. The MIT Press.

Koivo, H. (2001). *Soft Computing in Dynamical Systems*. Helsinki University of Technology.

Liebemann, E., & Meder, K. (2004). Safety and performance enhancement: the Bosch electronic stability control (ESP). *SAE Paper*.

Lincke, W., Richter, B., & Schmidt, R. (1973). Simulation and measurement of driver vehicle handling performance. *SAE Paper*, *730489*.

Macadam, C. C. (2003). Understanding and Modeling the Human Driver. *Vehicle System Dynamics*, *40*(1-3), 101–134. doi:10.1076/vesd.40.1.101.15875

MacDonald, & Hoffmann. (1980). Review of relationships between steering wheel reversal rate and driving task demand. doi:10.1177/001872088002200609

Martens, M., & Winsum, W. Van. (2000). Measuring distraction: the peripheral detection task. *TNO Human Factors, ….*

MathWorks. (2014). *Optimization Toolbox™ User's Guide* (2014b ed.). The MathWorks, Inc.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, *5*(4), 115–133.

McRuer, D. T., Graham, D., Krendel, E., & Reisener Jr, W. (1965). Human pilot dynamics in compensatory systems. SYSTEMS TECHNOLOGY INC HAWTHORNE CA.

McRuer, D. T., & Krendel, E. S. (1962). The Man-Machine System Concept. *Proceedings of the IRE*, *50*(5), 1117–1123.

McRuer, D. T., & Weir, D. H. (1969). Theory of Manual Vehicular Control. *Ergonomics*, *12*(4), 599–633.

McRuer, & Krendel. (1959a). The human operator as a servo system element Part I.

McRuer, & Krendel. (1959b). The human operator as a servo system element Part II.

Megaw, E. (2005). The definition and measurement of mental workload. *Evaluation of Human Work*, 525–551.

Monsma, S., & Arts, L. (2008). Tire characteristics and their influence on driver models. *Tire Technology International, Annual Review*.

Monsma, S., Buning, L., Pauwelussen, J., & Bremmer, P. (2008). The Human Touch in Tyre Handling Performance Assessment, a Model Based Approach. *AVEC*.

Monsma, S., & Defilippi, E. (2011). Artificial neural Network for the Assessment of Driver Judgement (p. IAVSD 2011).

Monsma, S., & Oort, van, M. J. M. (2010). Subjective evaluation of handling behaviour related to tyre dependent driver parameters. In *AVEC* (pp. 574–597). Loughborough, UK.

Monsma, S., & Sultania, S. (2011). Quantification of Drivers Mental Workload During Outdoor Testing Using Heart Rate Variability. *EAEC*.

Mulder, L. J. M. (1992). Measurement and analysis methods of heart rate and respiration for use in applied environments. *Biological Psychology*, *34*(2-3), 205–236.

Munakata, T. (1998). Fundamentals of the new artificial intelligence: Beyond traditional paradigms. Springer (New York).

O'Donnell, R. D., & Eggemeier, F. T. (1986). Handbook of Perception and Human Performance. Volume 2. Cognitive Processes and Performance. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of Perception and Human Performance. Volume 2. Cognitive Processes and Performance* (Vol. II). New York, NY: John Wiley and Sons.

Ooki, M., Sakai, H., Muragishi, Y., Fukui, K., & Ono, E. (2008). Vehicle Transient Response Based on Human Sensitivity. *AVEC*.

Oort, van, M. J. M. (2011). Modelling steering behaviour; Development of a parameter-estimation method used for tyre-dependent driver model parameters.

Oort, van, M. J. M. (2012). *Evaluation of the Path Preview Error Driver Model Used for Predicting Mental Workload*. Eindhoven University of Technology.

Pacejka, H. B. (2012). *Tire and Vehicle Dynamics*. Butterworth-Heinemann.

Pauwelussen, J. P. (1999a). Effect of tyre handling characteristics on driver judgement of vehicle directional stability. *Vehicle Performance: Understanding Human Monitoring and Assessment*, 151.

Pauwelussen, J. P. (1999b). *Vehicle Performance: Understanding Human Monitoring and Assessment*. Taylor \& Francis Group.

Pauwelussen, J. P. (2012). Dependencies of driver steering control parameters. *Vehicle System Dynamics*, *50*(6), 939–959. doi:10.1080/00423114.2011.651476

Pauwelussen, J. P. (2014). *Essentials of Vehicle Dynamics* (1st ed.). Butterworth-Heinemann, an imprint of Elsevier SAE edition.

Pauwelussen, J. P., Dalhuijsen, W., & Merts, M. (2007). *Tyre dynamics, tyre as a vehicle component Part 1.: Tyre handling performance*.

Pauwelussen, J. P., & Pauwelussen, J. J. A. (2004). Exploration of Steering Wheel Angle Based Workload Measures in Relationship to Steering Feel Evaluation. *AVEC*.

Plöchl, M., & Edelmann, J. (2007). Driver models in automobile dynamics application. *Vehicle System Dynamics*, *45*(7), 699–741.

Prokhorov, D. (2008). Neural Networks in Automotive Applications. In *Computational Intelligence in Automotive Applications* (Vol. 132, pp. 101–123). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-79257-4

Prokop, G. (2001). Modeling Human Vehicle Driving by Model Predictive Online Optimization. *Vehicle System Dynamics*, *35*(1), 19–53.

Pronk, G. M., Stassen, H. G., Lunteren Ton, V., Hosman, R. J. A. W., Vaart Hans C., van der, & Helm Frans C.T., van der. (2000). *Cybernetische Ergonomie*.

Reid, L. D. (1983). A survey of recent driver steering behavior models suited to accident studies. *Accident Analysis \& Prevention*, *15*(1), 23–40.

Reid, L. D., Solowka, E. N., & Billing, A. M. (1981). A systematic study of driver steering behaviour. *Ergonomics*, *24*(6), 447–462.

Renault. (2014). Drive the change. Retrieved December 1, 2014, from http://www.renault.com/en/pages/home.aspx

Renner, T. E., & Barber, A. J. (2000). Accurate Tire Models for Vehicle Handling Using the Empirical Dynamics Method. In *International ADAMS User Conference*.

Russell, S. (2009). Artificial Intelligence: A Modern Approach Author: Stuart Russell, Peter Norvig, Publisher: Prentice Hall Pa.

Salvucci, D. D., & Gray, R. (2004). A two-point visual control model of steering. *Perception*, *33*, 1233–1248.

Savkoor, A. R., Happel, H., & Horkay, F. (1999). Vehicle handling and sensitivity in transient manoeuvres. *Vehicle Performance: Understanding Human Monitoring and Assessment*, 121–147.

Seghers, G. (2006). Study to improve "objective-subjective" correlation techniques in the framework of vehicle dynamics.

Sharp, R. S. (1999). Vehicle Dynamics and the Judgement of Quality. *Vehicle Performance: Understanding Human Monitoring and Assessment*, 87–96.

Sharp, R. S., Casanova, D., & Symonds, P. (2000). A Mathematical Model for Driver Steering Control , with Design, Tuning and Performance results. *Vehicle System Dynamics*, (768414408). doi:10.1076/0042-3114(200005)33

Specht, D. F. (1991). A general regression neural network. *Neural Networks, IEEE Transactions on*, *2*(6), 568–576. doi:10.1109/72.97934

Stassen, Johanssen, & Moray. (1990). Internal representation, internal model, human performance model and mental workload.

Toffin, D., Reymond, G., Kemeny, A., & Droulez, J. (2007). Role of steering wheel feedback on driver performance: driving simulator and modeling analysis. *Vehicle System Dynamics*, *45*(4), 375–388. doi:10.1080/00423110601058874

Tseng, H. E., Asgari, J., Hrovat, D., van der Jagt, P., Cherry, A., & Neads, S. (2005). Evasive manoeuvres with a steering robot. *Vehicle System Dynamics*, *43*(3), 199–216. doi:10.1080/00423110420002667775

Tustin, A. (1947). The nature of the operator's response in manual control, and its implications for controller design. *Journal of the Institution of Electrical Engineers - Part IIA: Automatic Regulators and Servo Mechanisms*, *94*(2), 190–206. doi:10.1049/ji-2a.1947.0025

Van Elslande, P., & Fouquet, K. (2007). Analyzing'human functional failures' in road accidents. *Final Report. Deliverable D*, *5*.

Van Randwijk, M. J., Godthelp, J., Käppler, W. D., & Ruys, P. A. J. (1991). 'Correlation of driver judgements and vehicle directional data to evaluate and predict truck handling'. *TNO, FGAN*.

Verwey, W. B., & Veltman, H. A. (1996). Detecting short periods of elevated workload: A comparison of nine workload assessment techniques. *Journal of Experimental Psychology: Applied*, *2*(3), 270–285.

Volvo Cars. (2014). Volvo. For Life. Retrieved December 1, 2014, from http://www.volvocars.com/intl/campaigns/introduction/philosophy/pages/default.aspx

Von Glasner, E. G., & Ahlgrimm, J. (2011). The influence of tyre characteristics on the dynamic performance of vehicles. In *13th European Automotive Congress, EAEC*. Valencia, Spain.

Vos, de, A. P., Godthelp, J., & Käppler, W. D. (1999). Subjective and objective assessment of manual, supported, and automated vehicle control. *Vehicle Performance: Understanding Human Monitoring and Assessment*, 97–120.

Waard, de, D. (1996). *The measurement of drivers' mental workload. Traffic Research Centre VSC*. University of Groningen.

Waard, de, D., Kruizinga, A., & Brookhuis, K. A. (2008). The consequences of an increase in heavy goods vehicles for passenger car drivers' mental workload and behaviour: A simulator study. *Accident Analysis \& Prevention*, *40*(2), 818–828.

Weir, D. H., & Di Marco, R. J. (1978). Correlation and evaluation of driver/vehicle directional handling data. *Sae Journal*.

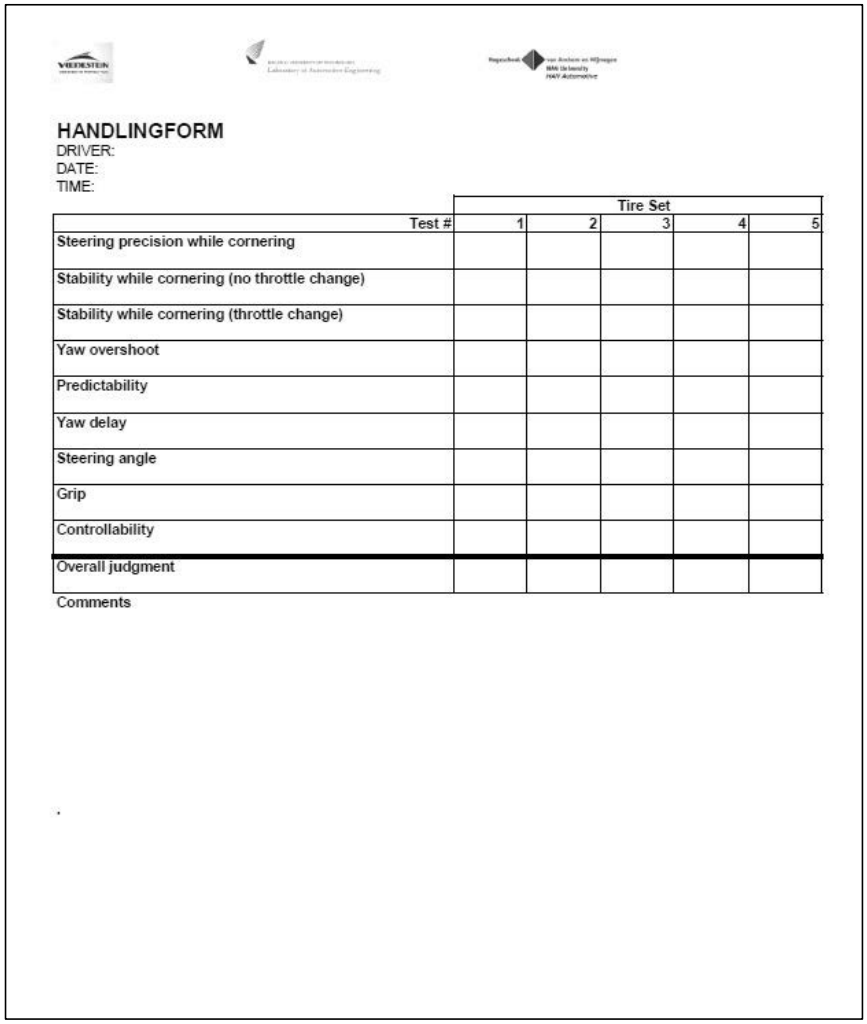
Wickens, C. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors ...*.

Yuhara, N., & Tajima, J. (2000). Advanced Steering System Adaptable to Lateral Control Task and Driver's Intention. *AVEC*.

Zijlstra, F., & Doorn, L. Van. (1985). *The construction of a scale to measure perceived effort*.

Appendix 1 Subjective Assessment Form

Fig. 128 shows the subjective assessment form (English translation) including the 10 handling aspects to be scored by the driver for different tire sets.



HANDLINGFORM

DRIVER:
DATE:
TIME:

	Tire Set				
Test #	1	2	3	4	5
Steering precision while cornering					
Stability while cornering (no throttle change)					
Stability while cornering (throttle change)					
Yaw overshoot					
Predictability					
Yaw delay					
Steering angle					
Grip					
Controllability					
Overall judgment					

Comments

Figure 128. Subjective assessment form.

Appendix 2 Rating Scale Mental Effort

Fig. 129 shows the Rating Scale Mental Effort (Zijlstra & Doorn, 1985) (English translation).

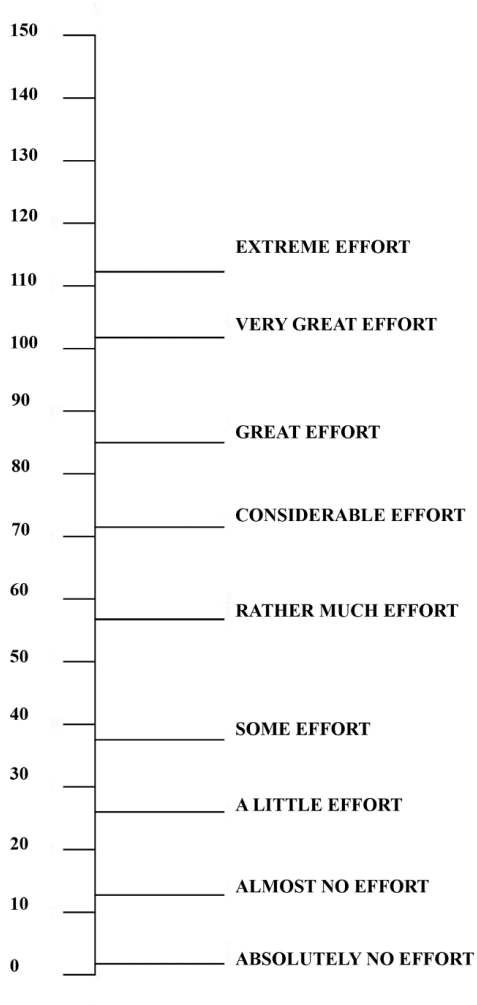


Figure 129. Rating Scale Mental Effort.

Appendix 3 Signal Parts and Metric Calculations

In this appendix the definition of the signals parts are given, which are used for calculating the metrics that can be used as predictors for the GRNN, as explained in section 6.1.1. For each signal, an example time plot of the double lane change is shown for tire 1 and tire 6 to show the distinguishing signal behavior between two different tires driven by the same driver. Note that the definition of the signal parts is based on all driven lane changes for all tires for both drivers, therefore it is possible that the signal definition for this specific example does not seem to be the most distinguishing part.

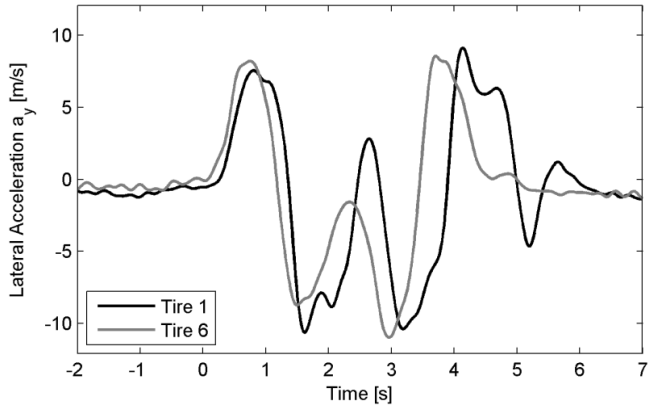


Figure 130. Double lane change time plot of the lateral acceleration a_y.

Table 18. Signal parts of the lateral acceleration a_y.

Signal Part Lateral Acceleration	Time t [s]	Metric Calculation
I	-2.0 - 1.0	int (neg)
II	-0.5 – 2.0	int (pos)
III	1.0 – 4.0	int (neg)
IV	3.0 – 6.0	int (pos)
V	4.0 – 7.0	int (neg)

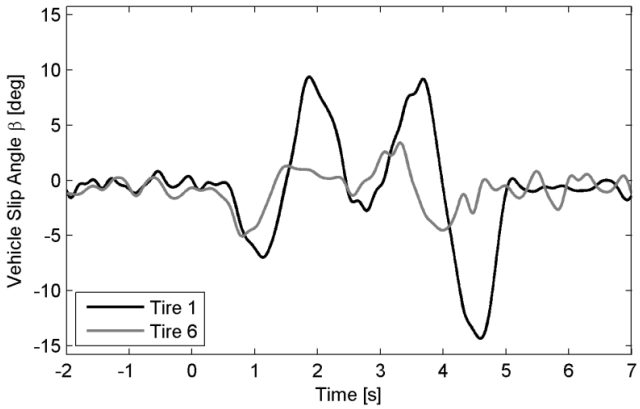


Figure 131. Double lane change time plot of the vehicle slip angle β.

Table 19. Signal parts of the vehicle slip angle β.

Signal Part Vehicle Slip Angle	Time t [s]	Metric Calculation
I	0.0 - 2.0	int (neg)
II	1.5 – 3.5	int (neg)
III	4.0 – 5.0	int (neg)
IV	0.0 – 2.0	min

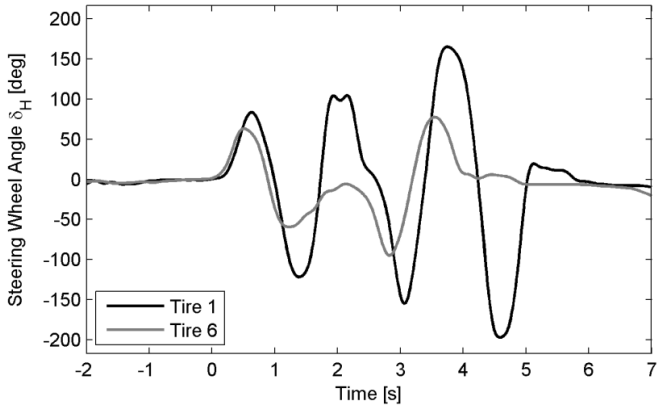


Figure 132. Double lane change time plot of the steering wheel angle δ_H .

Table 20. Signal parts of the steering wheel angle δ_H.

Signal Part Steering Wheel Angle	Time t [s]	Metric Calculation
I	0.0 - 1.5	max
II	1.0 – 2.0	min
III	1.5 – 3.0	int (pos)
IV	2.0 – 4.0	min
V	3.0 – 4.5	max

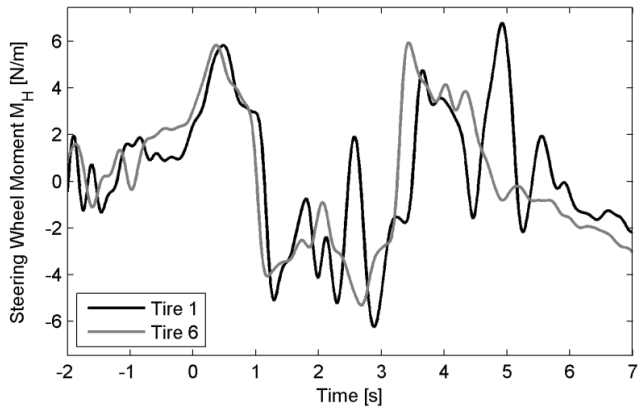


Figure 133. Double lane change time plot of the steering wheel moment M_H.

Table 21. Signal parts of the steering wheel moment M_H.

Signal Part **Steering Wheel Moment**	**Time t [s]**	**Metric Calculation**
I	3.0 - 4.0	int (pos)
II	1.0 – 2.0	int (neg)

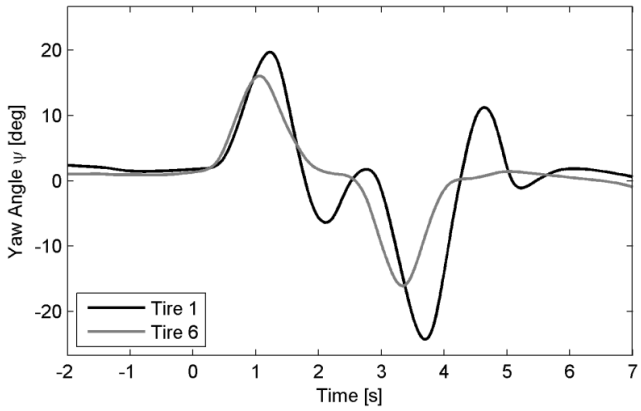


Figure 134. Double lane change time plot of the yaw angle ψ.

Table 22. Signal parts of the yaw angle ψ.

Signal Part **Yaw Angle**	**Time t [s]**	**Metric Calculation**
I	0.0 - 3.0	max
II	3.5 – 4.0	int (neg)
III	4.0 – 6.0	max

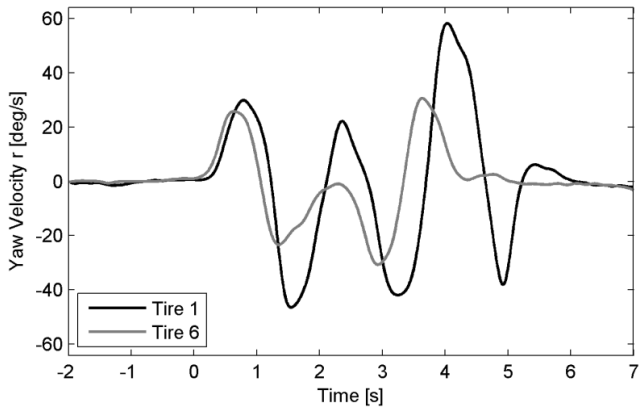


Figure 135. Double lane change time plot of the yaw velocity r.

Table 23. Signal parts of the yaw velocity r.

Signal Part Yaw Velocity	Time t [s]	Metric Calculation
I	1.0 - 2.0	int (neg)
II	3.0 – 4.0	int (pos)
III	4.0 – 5.0	int (neg)

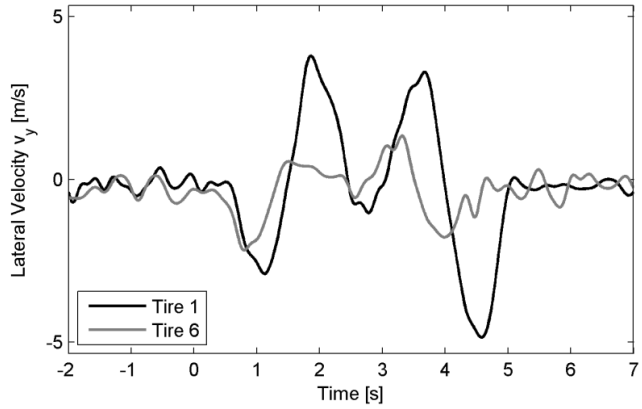


Figure 136. Double lane change time plot of the lateral velocity v_y.

Table 24. Signal parts of the lateral velocity v_y.

Signal Part Lateral Velocity	Time t [s]	Metric Calculation
I	0.5 - 1.5	min
II	1.0 – 3.0	max
III	2.5 – 4.0	max
IV	3.5 – 5.0	min
V	2.5 – 3.5	int (pos)
VI	3.5 – 4.5	int (neg)

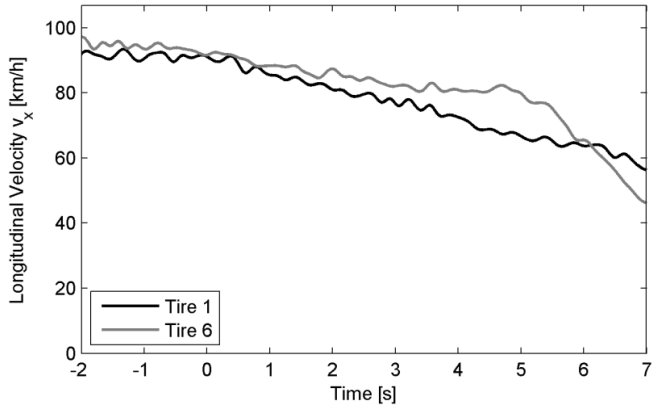


Figure 137. Double lane change time plot of the longitudinal velocity v_x.

Table 25. Signal parts of the longitudinal velocity v_x.

Signal Part Longitudinal Velocity	**Time t [s]**	**Metric Calculation**
I	-2.0 - 3.0	mean

Appendix 4 Metric Values Scatter Plots

Fig. 138 shows for every metric a separate scatter plot of the metric values for all subjective assessments for all drivers. As the differences between the metric values are of interest and not the absolute values, the numeric values are not shown on the *y*-axis. Fig. 139 shows the same plots for the standardized (zero mean, unity standard deviation) metrics.

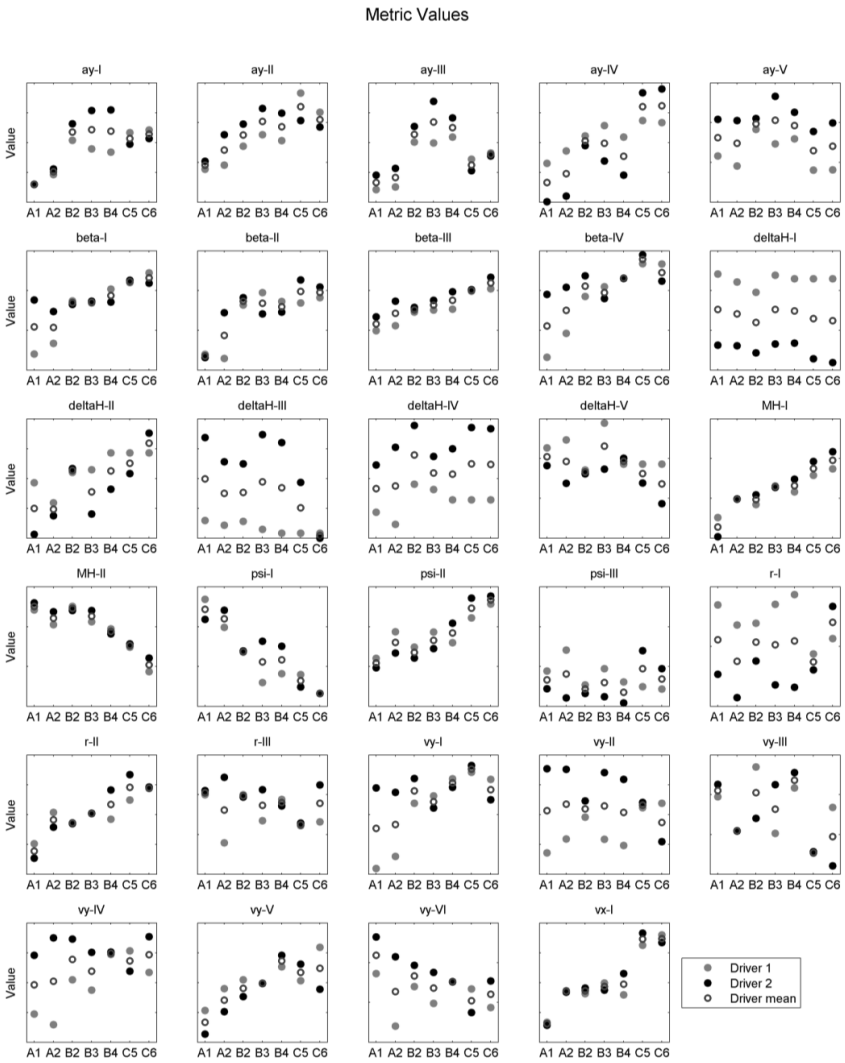


Figure 138. Scatter plots of metric values for all assessments for all drivers.

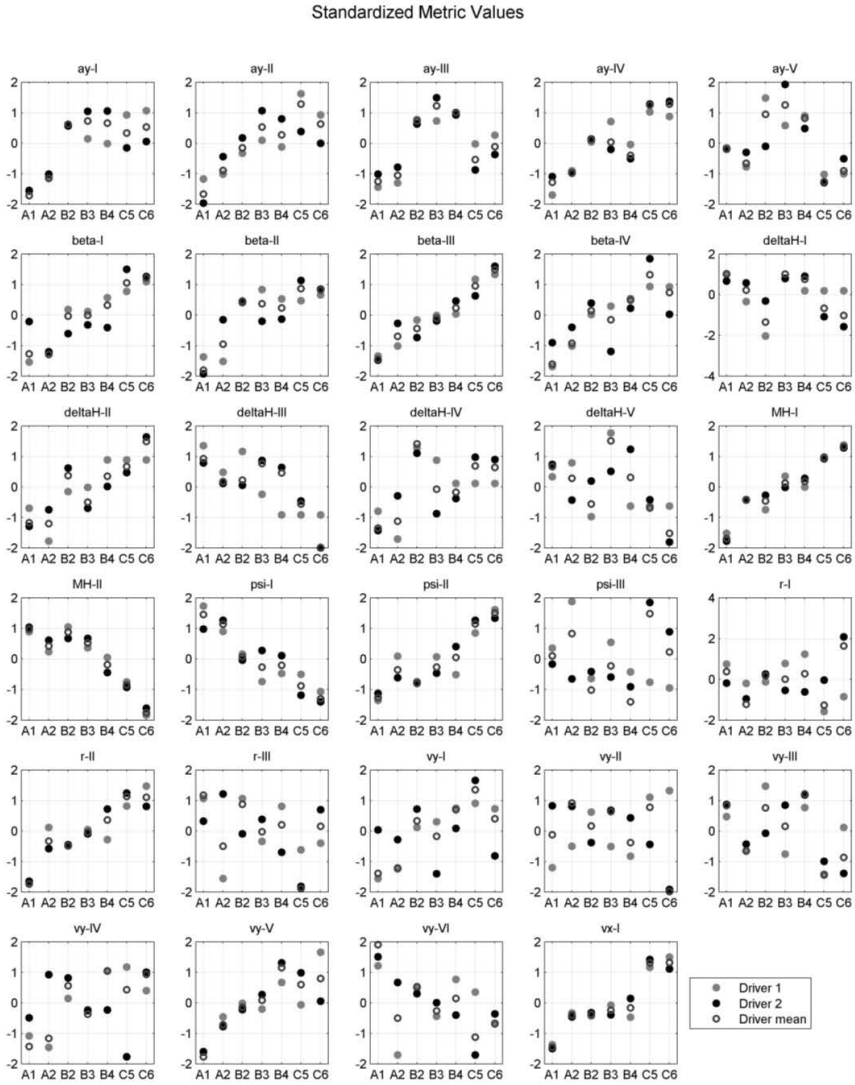


Figure 139. Scatter plots of standardized (zero mean, unity standard deviation) metric values for all assessments for all drivers.

Appendix 5 Subjective Assessments Scores

In this appendix the subjective assessments scores of the tire handling experiment, as described in section 4.6, are given. Each assessment is identified with a letter-number combination. The letter indicates the batch the tire is driven in: A, B or C; the number indicates the tire: 1 to 6. In Tables 26 to 28 the assessments scores are given for respectively driver 1, driver 2 and driver mean, the latter containing the mean value of the driver scores. Fig. 140 shows for every aspect a separate plot of the subjective assessments scores for all drivers.

Table 26. Driver 1 subjective assessments scores.

Assessment / Aspect	A1	A2	B2	B3	B4	C5	C6
1. Steering precision while cornering	5.00	6.00	6.00	6.75	7.25	7.00	8.00
2. Stability while cornering (no throttle change)	5.50	6.00	6.00	7.00	8.00	7.00	8.00
3. Stability while cornering (throttle change)	7.00	7.00	7.00	7.00	7.00	8.00	8.00
4. Yaw overshoot	7.00	7.00	8.00	8.00	8.00	8.00	8.00
5. Predictability	7.00	8.00	8.00	8.00	8.00	8.00	7.00
6. Yaw delay	7.00	7.00	7.00	7.00	7.00	7.00	7.00
7. Steering angle	5.50	6.00	6.00	6.50	7.00	7.00	8.00
8. Grip	5.50	6.00	6.00	6.50	7.00	7.00	8.00
9. Controllability	7.50	8.00	8.00	8.00	8.00	8.00	7.00
10. Overall judgement	**6.00**	**7.00**	**7.00**	**7.50**	**7.75**	**7.50**	**8.00**
Mean score	6,30	6,80	6,90	7,23	7,50	7,45	7,70

Highest score	8.00
Lowest score	5.00
Scoring range used	3.00

Table 27. Driver 2 subjective assessments scores.

Assessment / Aspect	A1	A2	B2	B3	B4	C5	C6
1. Steering precision while cornering	5.75	6.00	7.00	7.00	7.50	6.50	7.00
2. Stability while cornering (no throttle change)	7.00	7.00	7.00	7.00	7.50	7.00	7.00
3. Stability while cornering (throttle change)	6.50	7.00	7.00	7.00	7.50	6.50	6.50
4. Yaw overshoot	6.00	6.50	6.50	6.50	7.00	6.50	7.00
5. Predictability	7.00	7.00	7.00	7.00	7.50	7.00	7.00
6. Yaw delay	6.00	6.50	6.00	6.00	7.00	7.00	7.50
7. Steering angle	6.00	6.50	6.00	6.00	7.00	7.00	7.50
8. Grip	6.00	6.50	7.00	7.00	7.00	7.00	7.00
9. Controllability	7.00	7.00	7.00	7.00	7.00	7.00	7.00
10. Overall judgement	**6.00**	**6.50**	**6.50**	**6.50**	**7.50**	**7.00**	**7.50**
Mean score	6.33	6.65	6.70	6.70	7.25	6.85	7.10

Highest score	7.50
Lowest score	5.25
Scoring range used	2.25

Table 28. Driver mean subjective assessments scores.

Assessment / Aspect	A1	A2	B2	B3	B4	C5	C6
1. Steering precision while cornering	5.38	6.00	6.50	6.88	7.38	6.75	7.50
2. Stability while cornering (no throttle change)	6.25	6.50	6.50	7.00	7.75	7.00	7.50
3. Stability while cornering (throttle change)	6.75	7.00	7.00	7.00	7.25	7.25	7.25
4. Yaw overshoot	6.50	6.75	7.25	7.25	7.50	7.25	7.50
5. Predictability	7.00	7.50	7.50	7.50	7.75	7.50	7.00
6. Yaw delay	6.50	6.75	6.50	6.50	7.00	7.00	7.25
7. Steering angle	5.75	6.25	6.00	6.25	7.00	7.00	7.75
8. Grip	5.75	6.25	6.50	6.75	7.00	7.00	7.50
9. Controllability	7.25	7.50	7.50	7.50	7.50	7.50	7.00
10. Overall judgement	**6.00**	**6.75**	**6.75**	**7.00**	**7.63**	**7.25**	**7.75**
Mean score	6,31	6,73	6,80	6,96	7,38	7,15	7,40

Highest score	7.75
Lowest score	5.38
Scoring range used	2.37

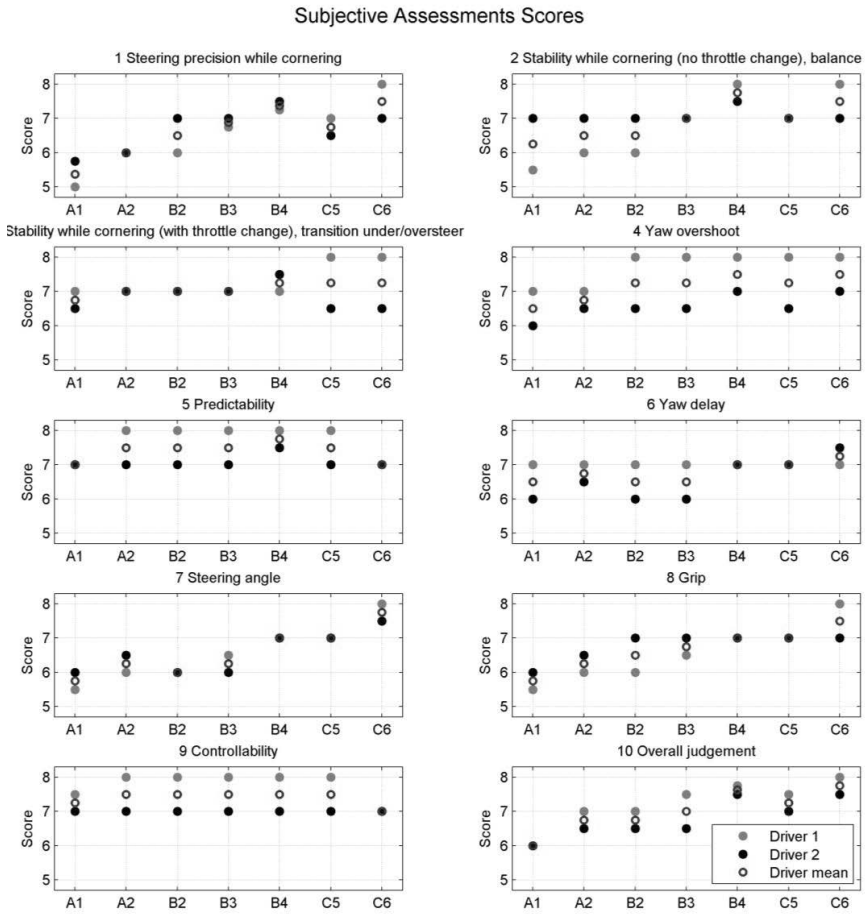


Figure 140. Plots of the subjective assessments scores per aspect for all drivers.

Appendix 6 Pearson's Correlation Coefficient Plots

For every driver and every aspect a figure is shown in this appendix including 29 plots; for every metric a scatter plot of metric – aspect values for all assessments. In addition, for every plot, the strength of the linear relationship is given by the value of Pearson's correlation coefficient r and a least squares line is plotted in the scatter plot. In every figure, the scatter plots are ordered from highest to lowest absolute r-value.

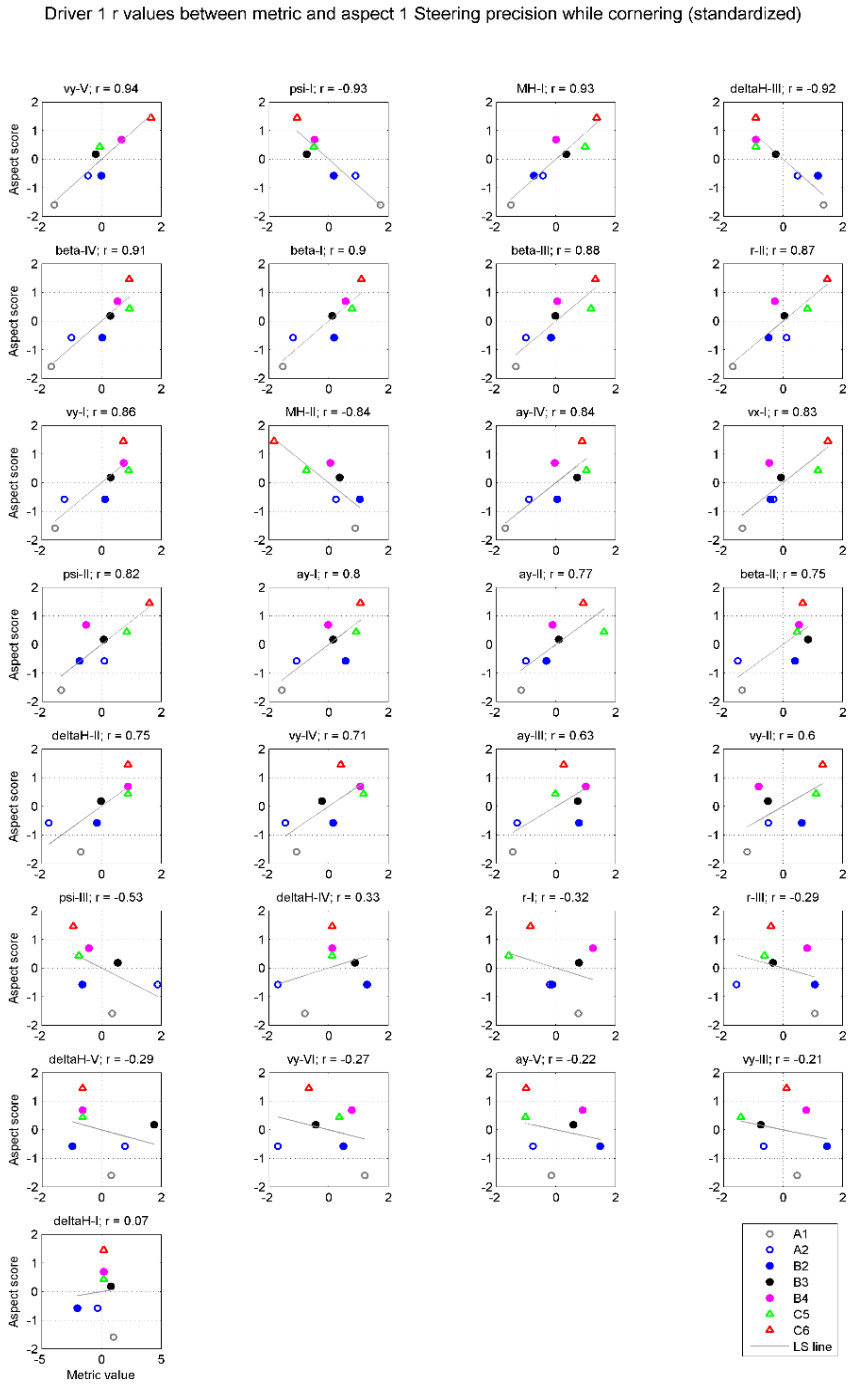


Figure 141. Driver 1 scatter plots, including the least squares line and r, of metric values and aspect scores 1, sorted from highest to lowest absolute r-value.

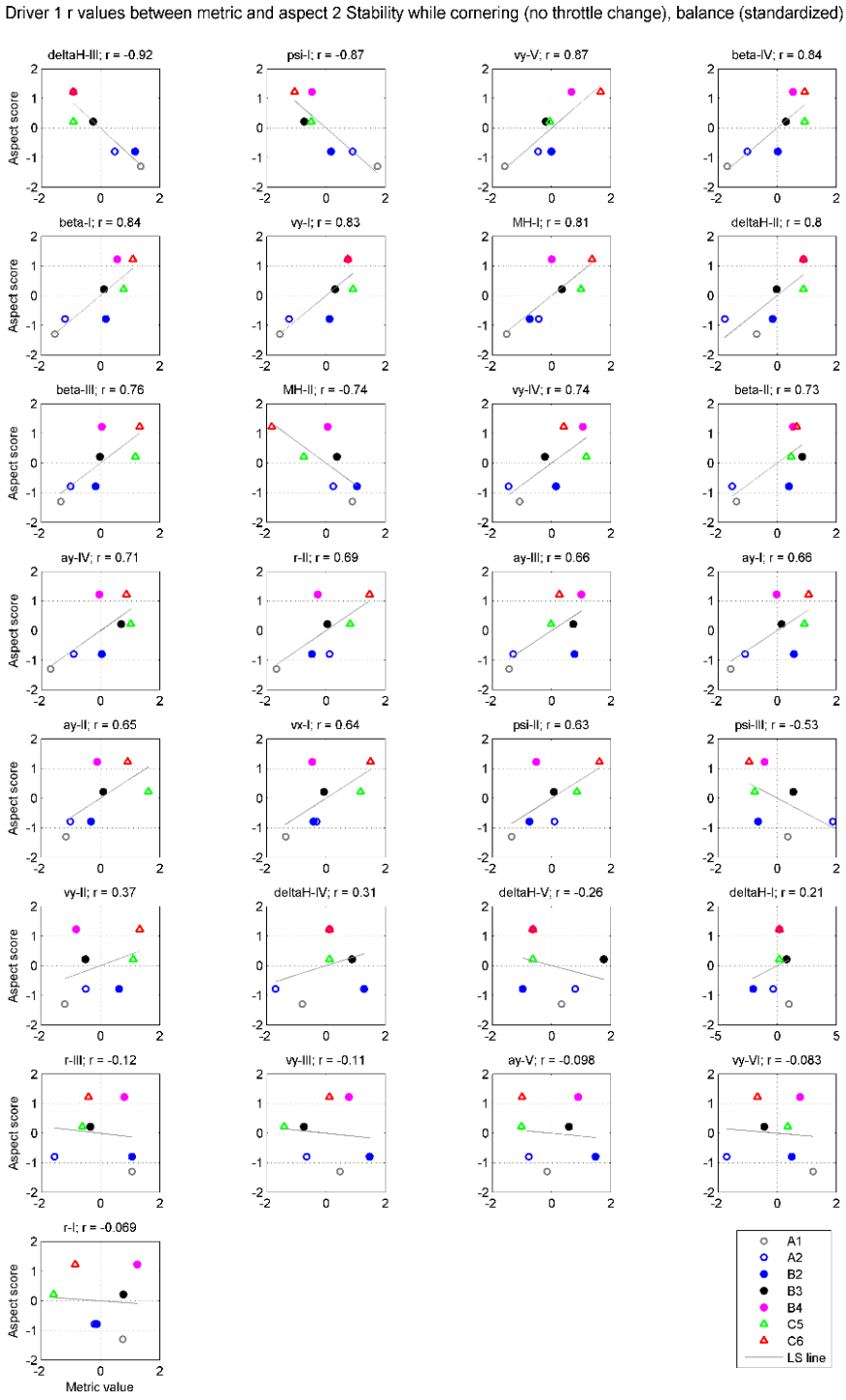


Figure 142. Driver 1 scatter plots, including the least squares line and r, of metric values and aspect scores 2, sorted from highest to lowest absolute r-value.

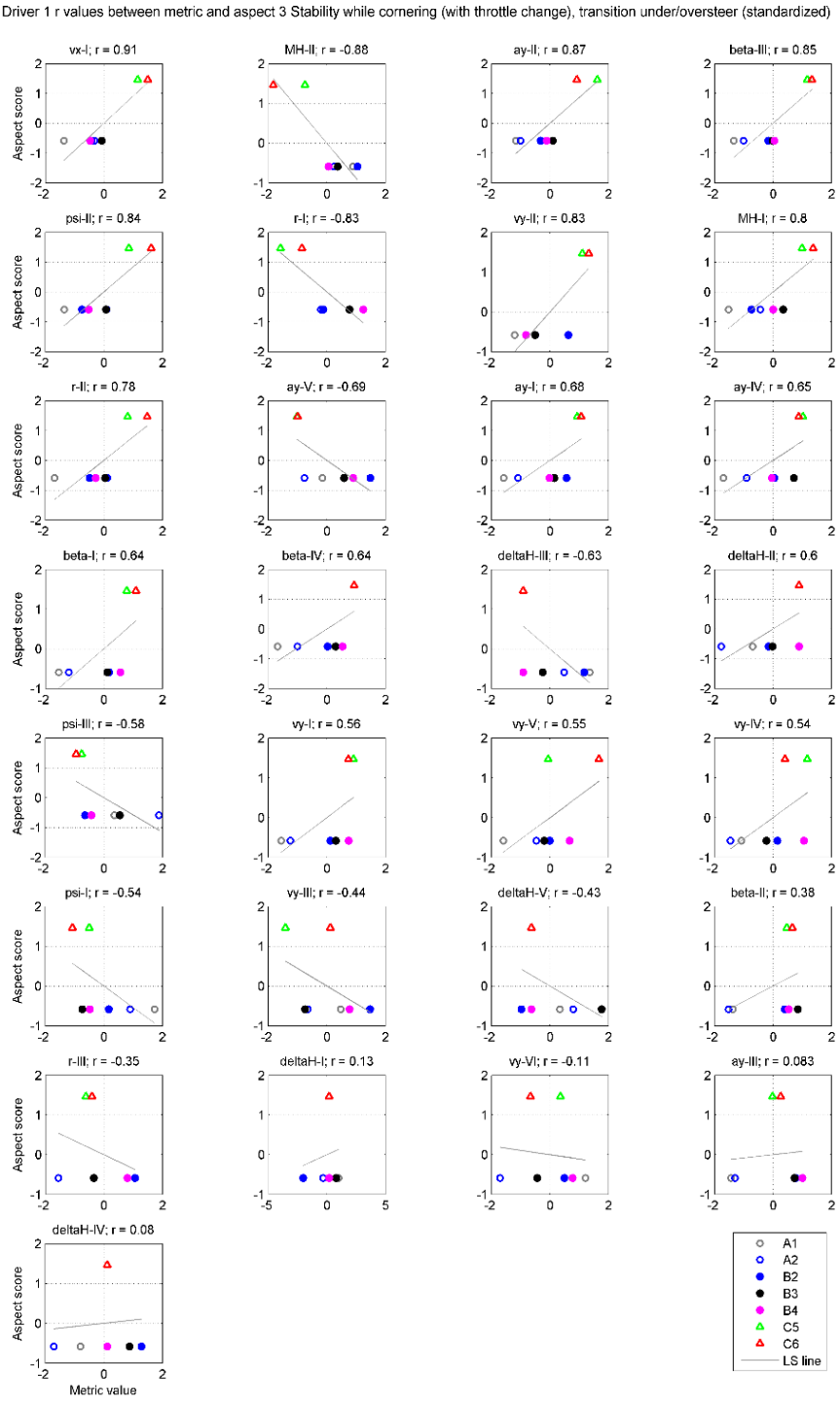


Figure 143. Driver 1 scatter plots, including the least squares line and r, of metric values and aspect scores 3, sorted from highest to lowest absolute r-value.

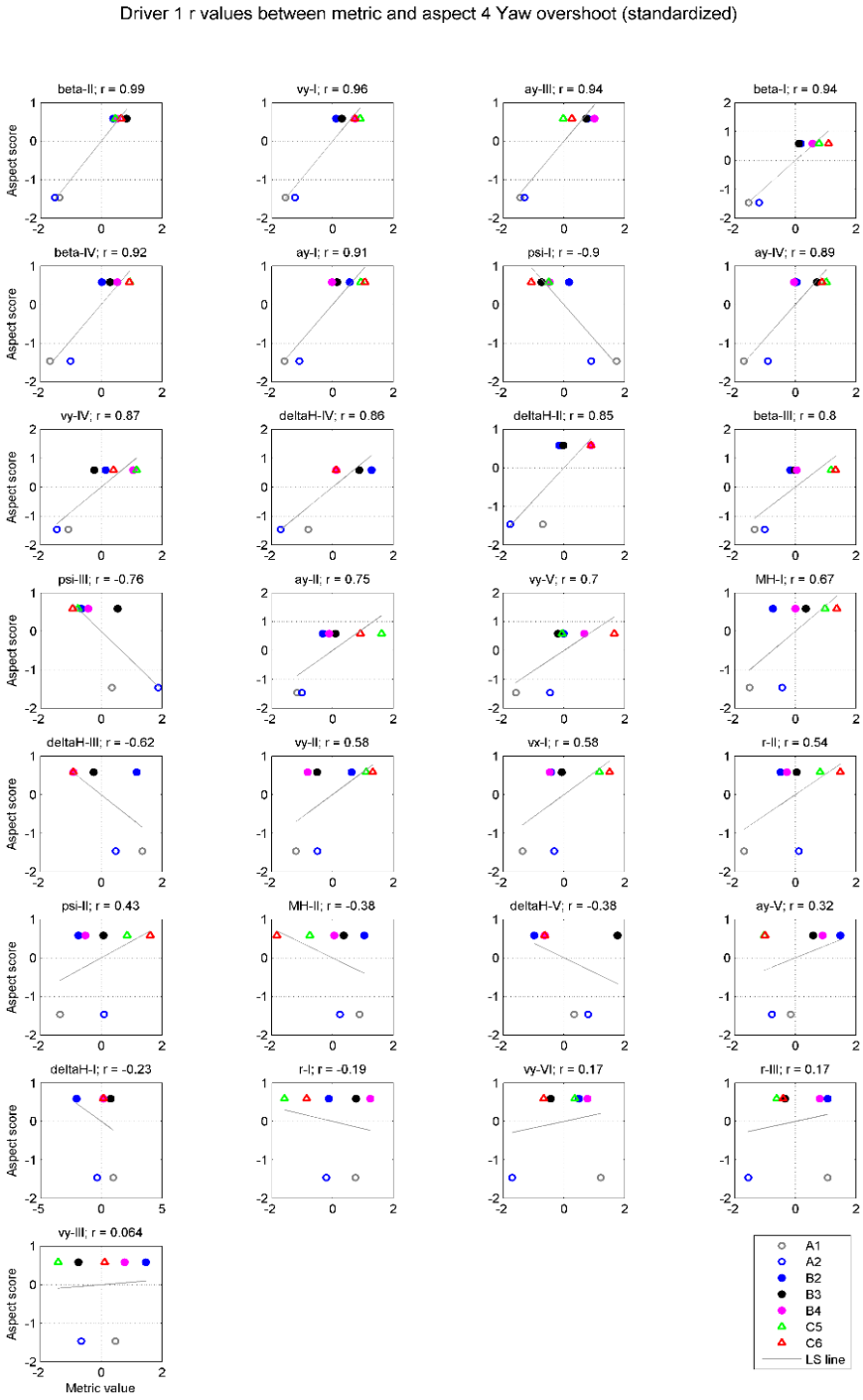


Figure 144. Driver 1 scatter plots, including the least squares line and r, of metric values and aspect scores 4, sorted from highest to lowest absolute r-value.

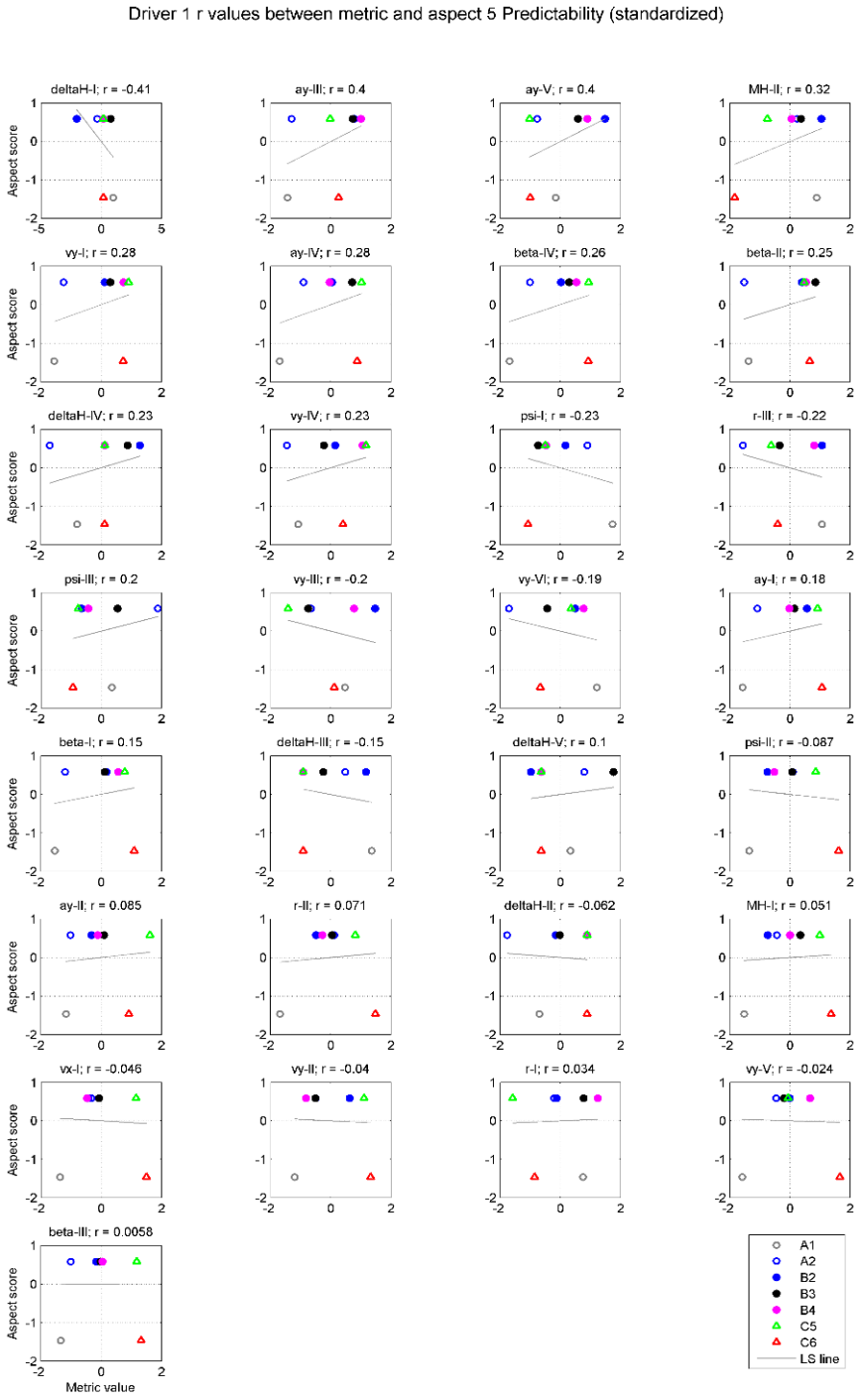


Figure 145. Driver 1 scatter plots, including the least squares line and r, of metric values and aspect scores 5, sorted from highest to lowest absolute r-value.

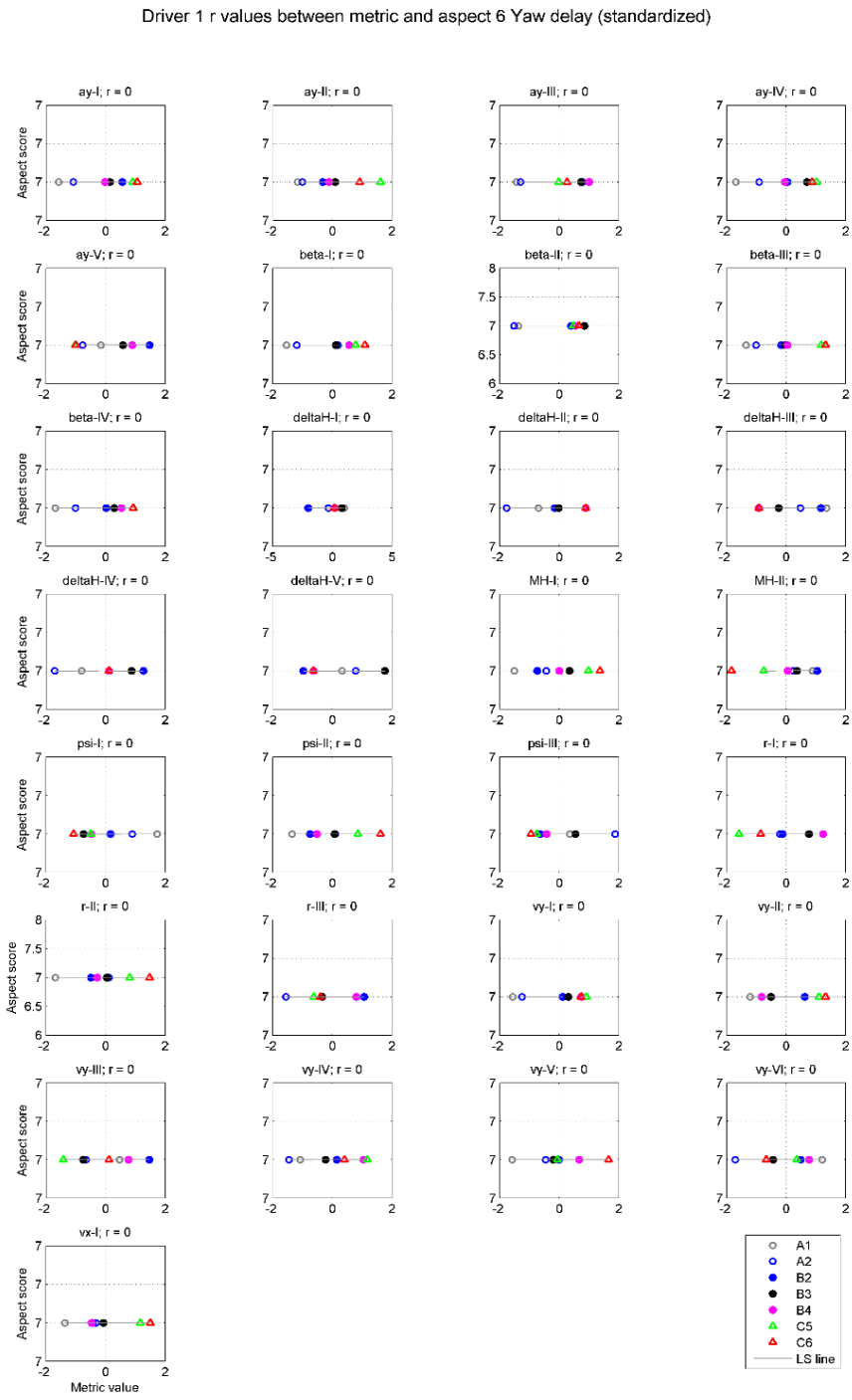


Figure 146. Driver 1 scatter plots, including the least squares line and r, of metric values and aspect scores 6, sorted from highest to lowest absolute r-value.

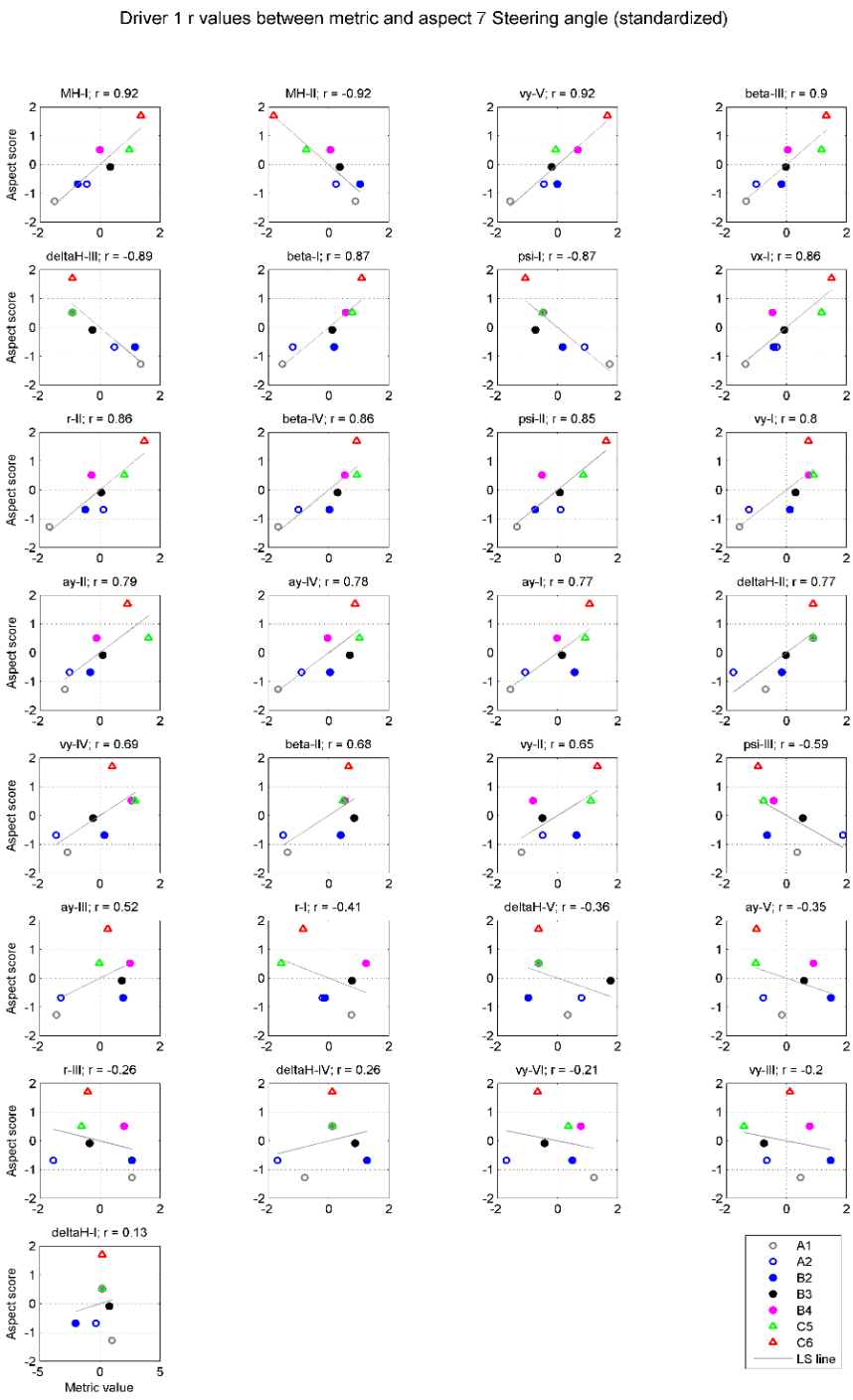


Figure 147. Driver 1 scatter plots, including the least squares line and r, of metric values and aspect scores 7, sorted from highest to lowest absolute r-value.

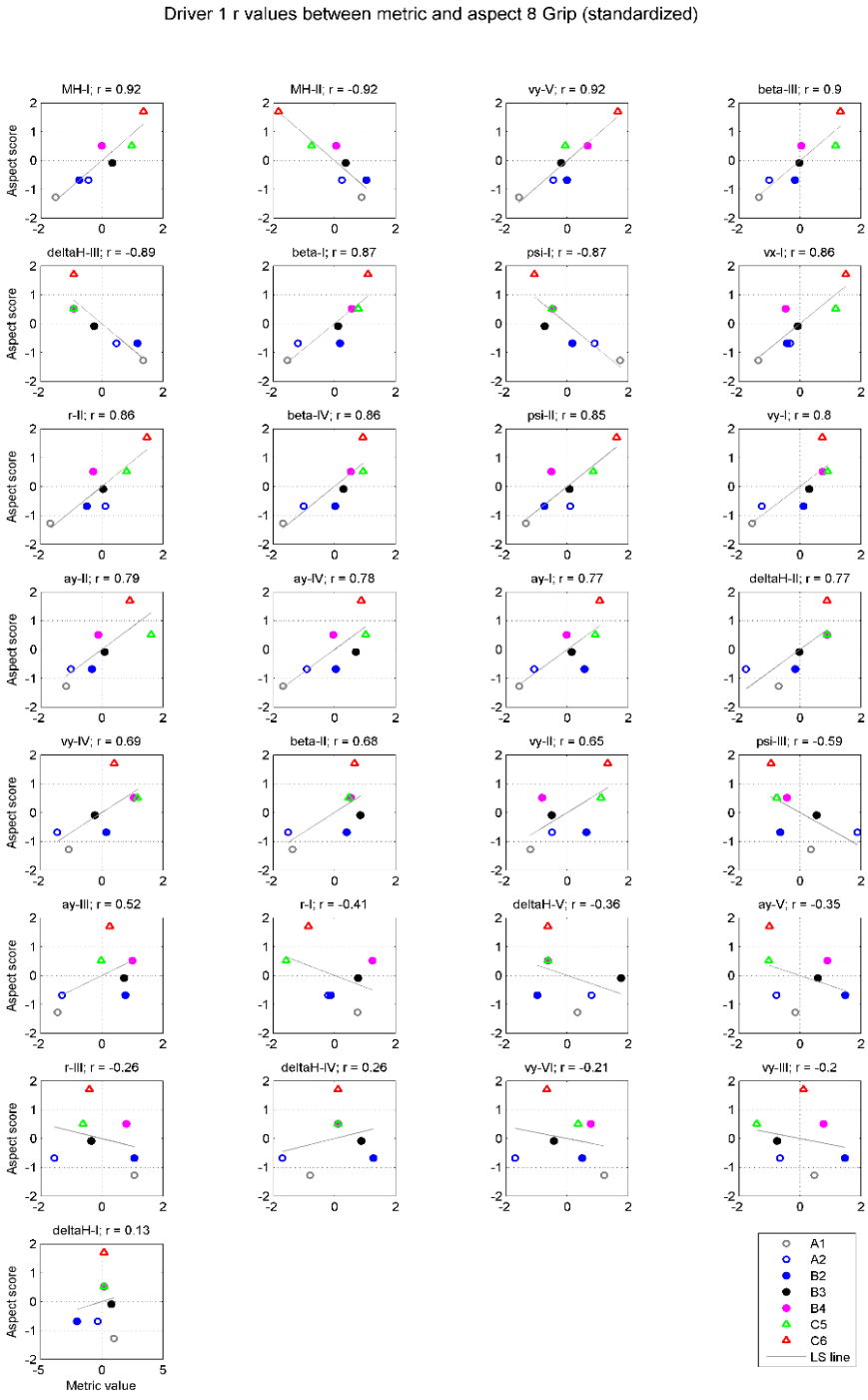


Figure 148. Driver 1 scatter plots, including the least squares line and r, of metric values and aspect scores 8, sorted from highest to lowest absolute r-value.

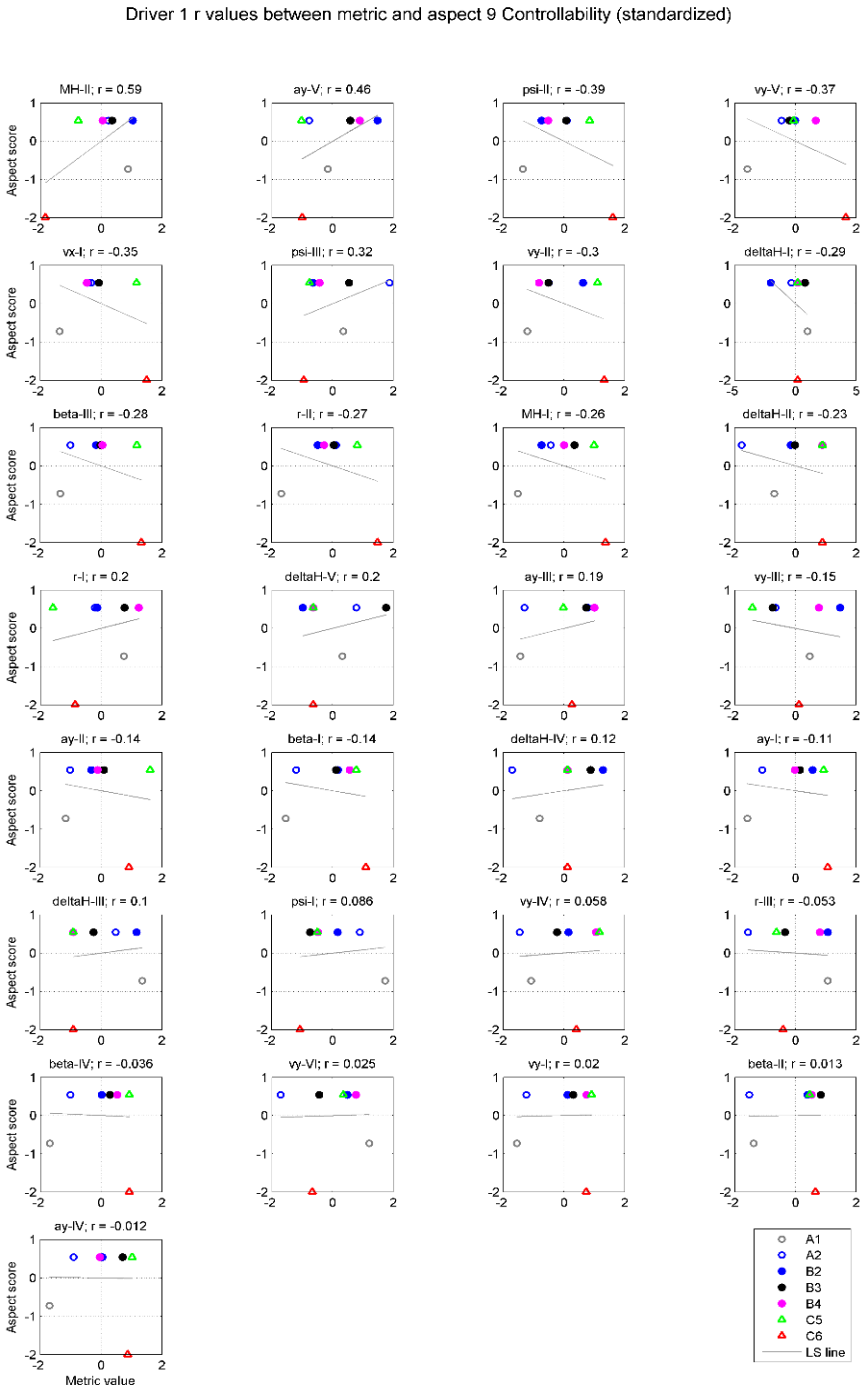


Figure 149. Driver 1 scatter plots, including the least squares line and r, of metric values and aspect scores 9, sorted from highest to lowest absolute r-value.

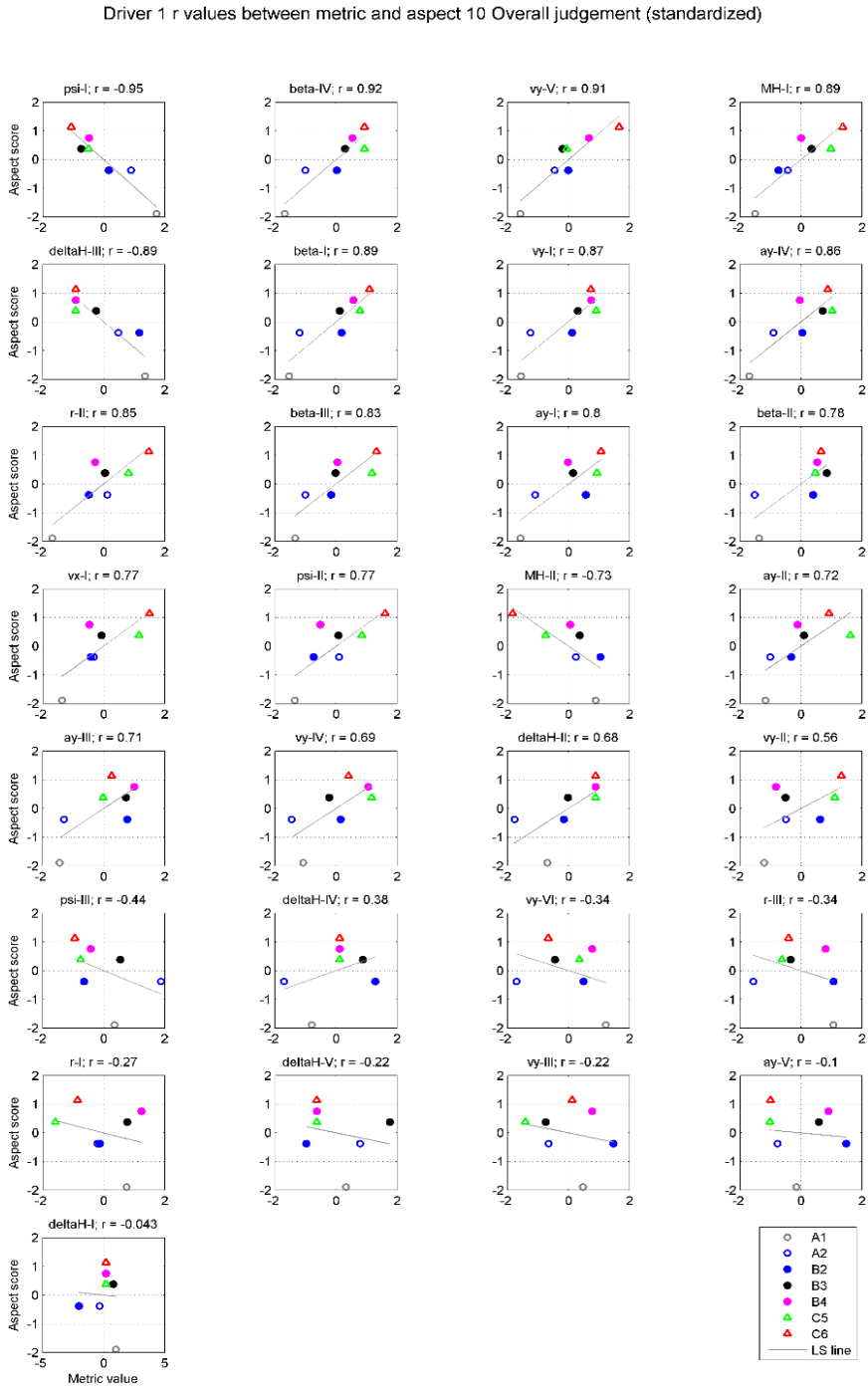


Figure 150. Driver 1 scatter plots, including the least squares line and r, of metric values and aspect scores 10, sorted from highest to lowest absolute r-value.

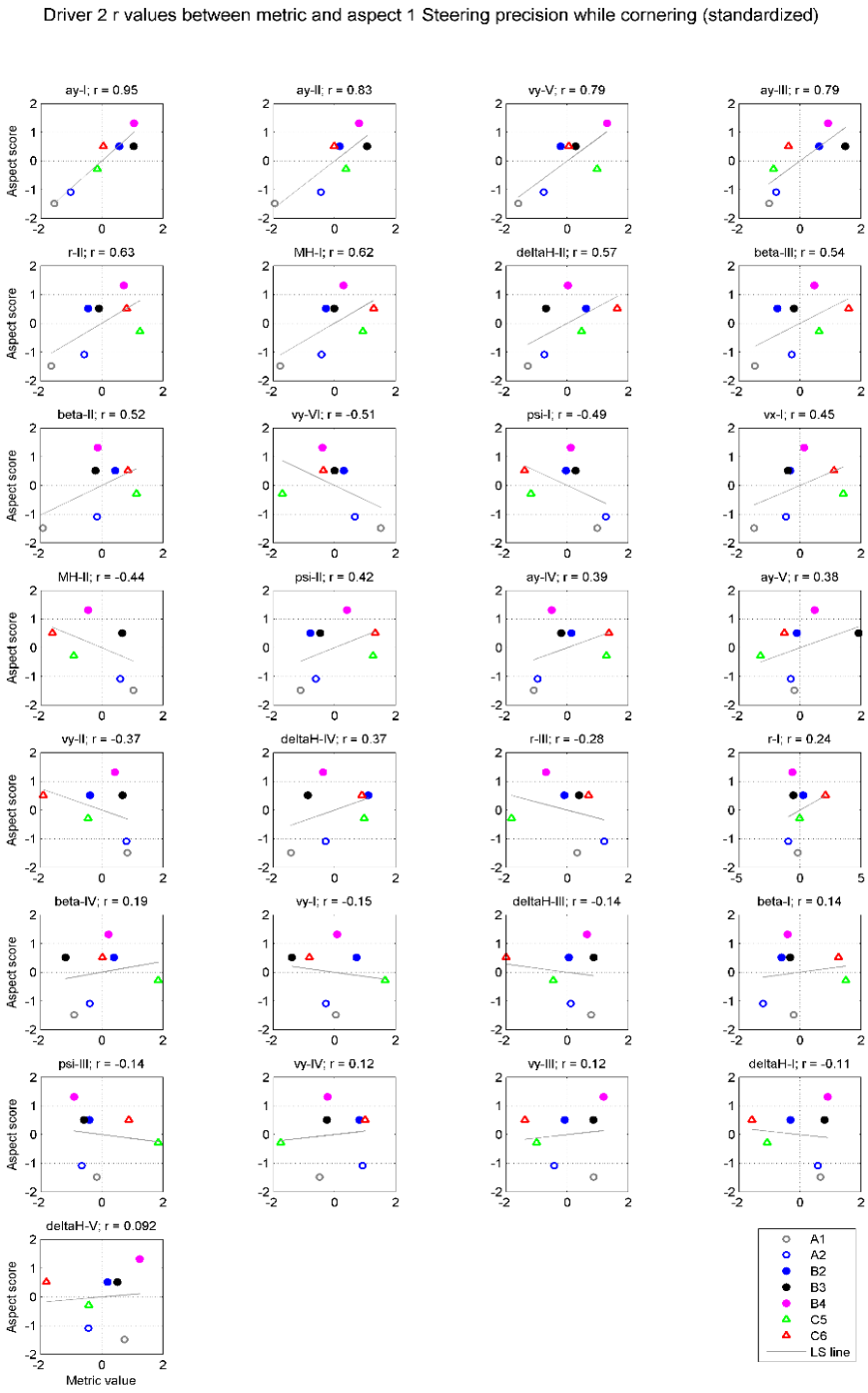


Figure 151. Driver 2 scatter plots, including the least squares line and r, of metric values and aspect scores 1, sorted from highest to lowest absolute r-value.

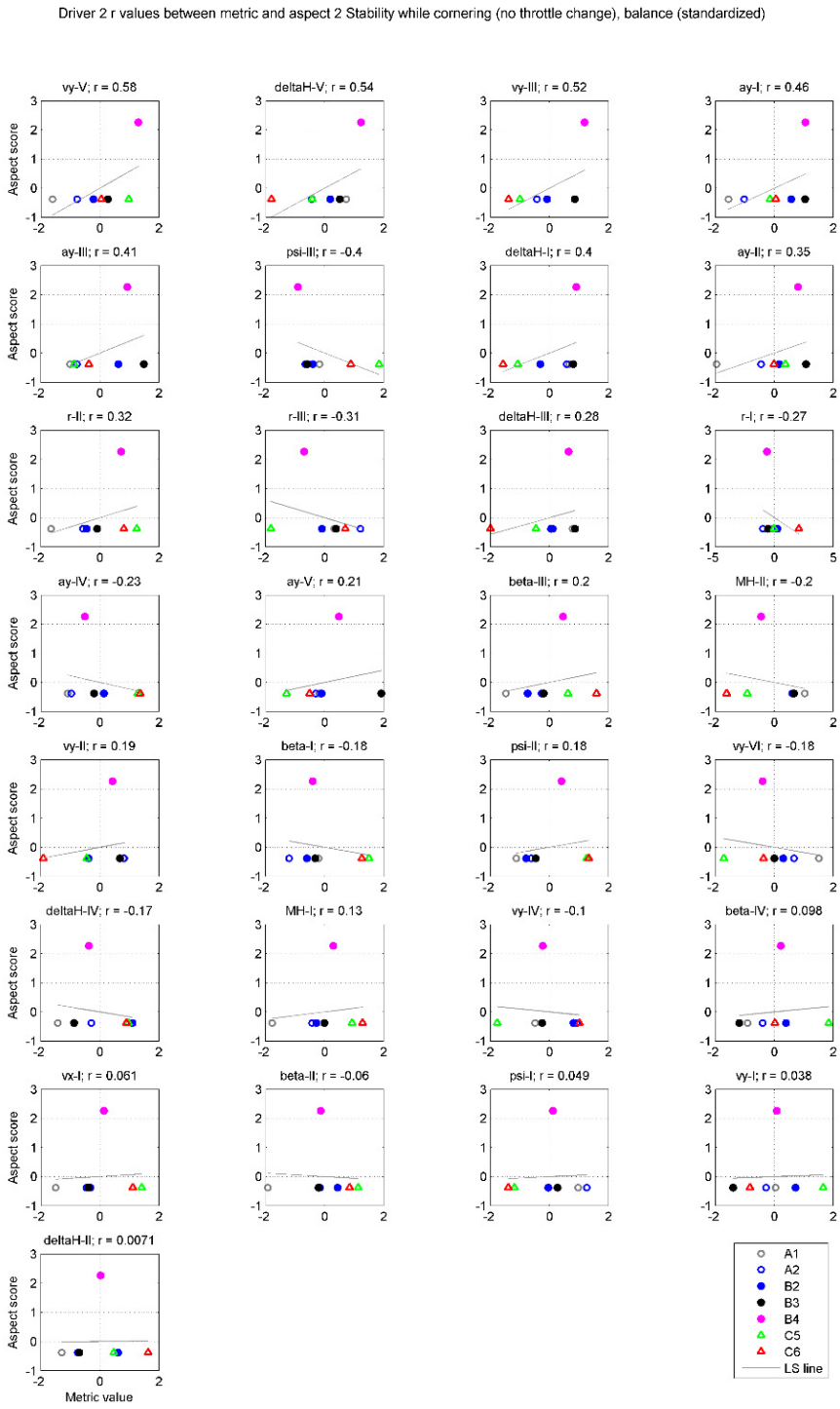


Figure 152. Driver 2 scatter plots, including the least squares line and r, of metric values and aspect scores 2, sorted from highest to lowest absolute r-value.

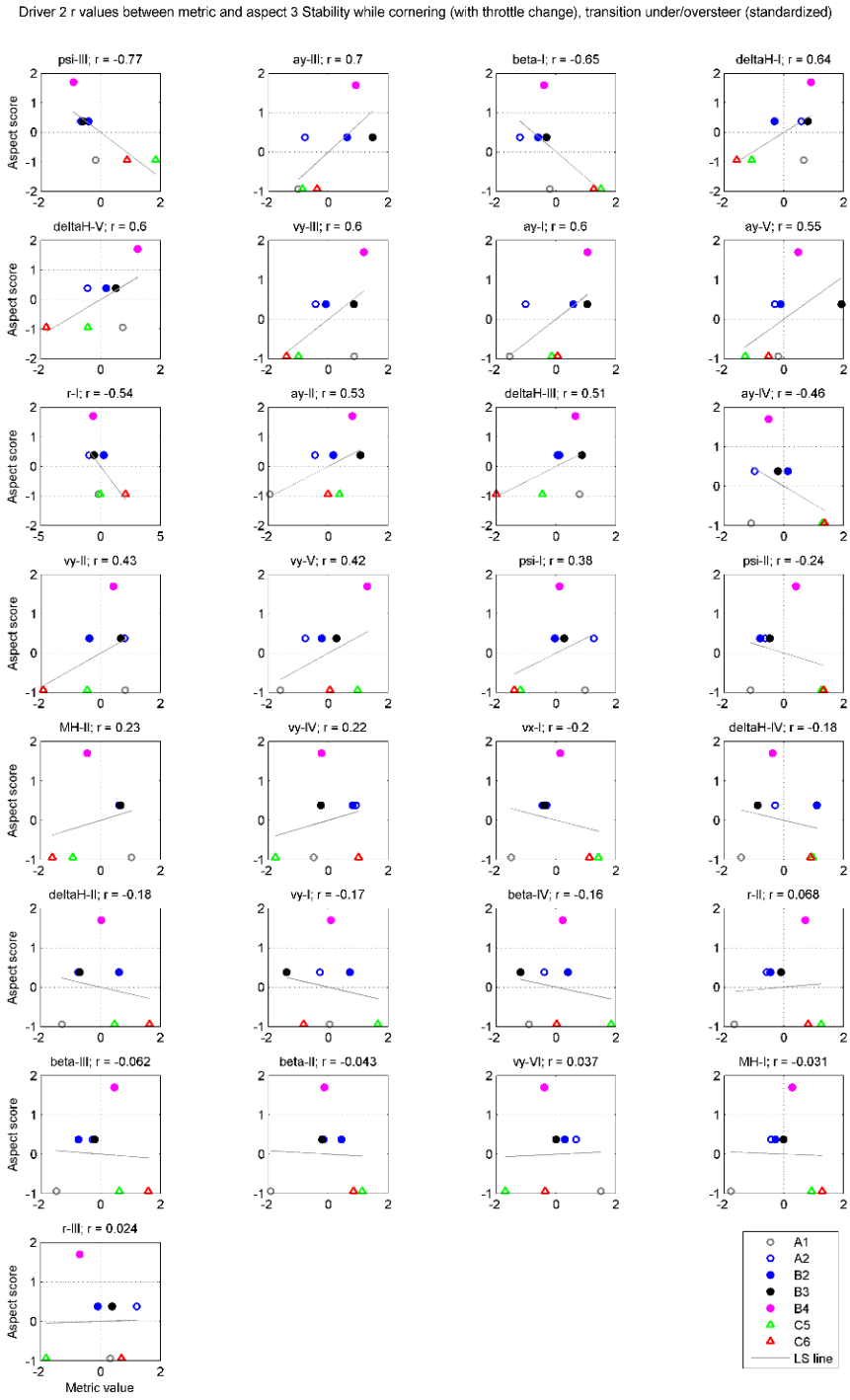


Figure 153. Driver 2 scatter plots, including the least squares line and r, of metric values and aspect scores 3, sorted from highest to lowest absolute r-value.

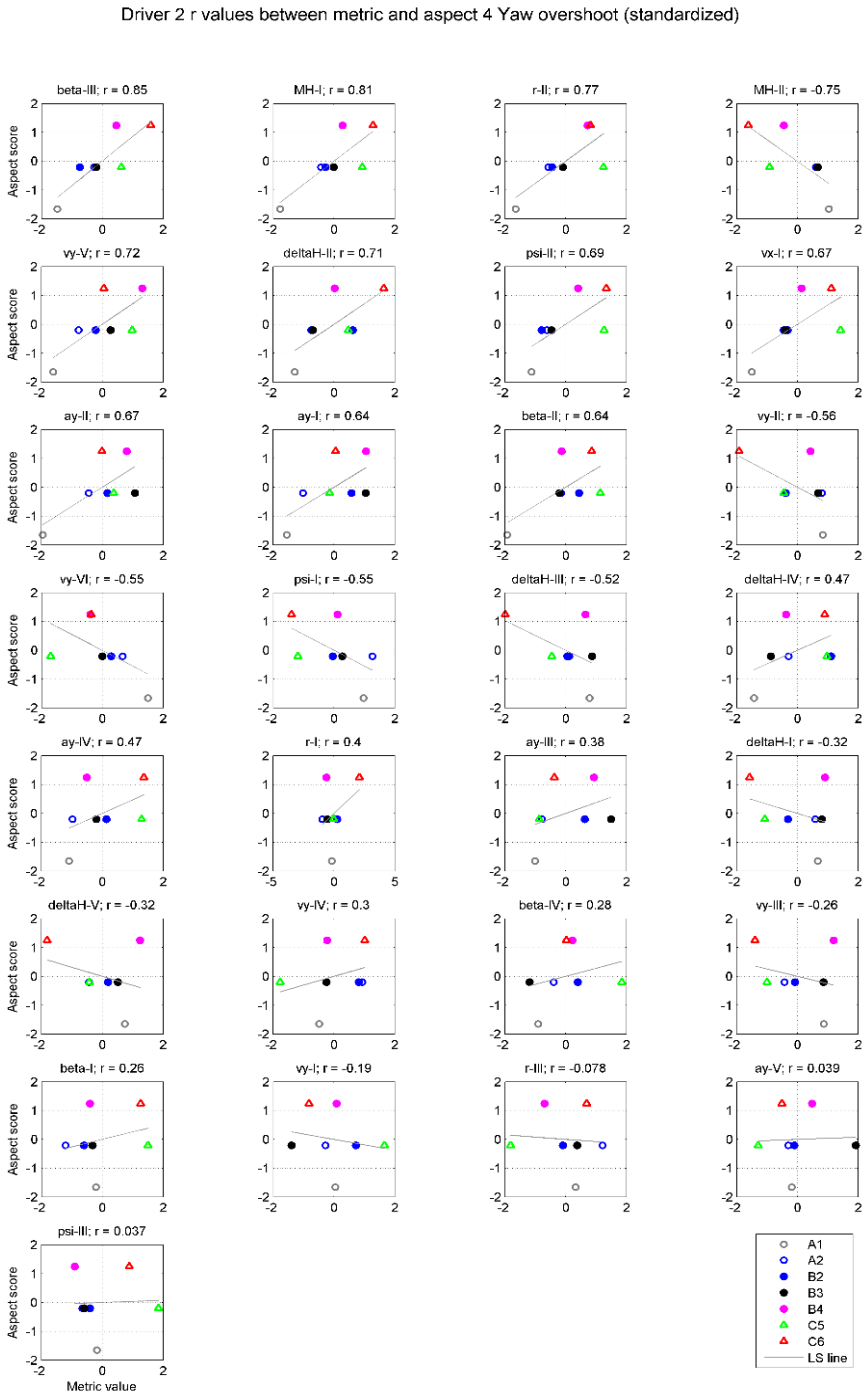


Figure 154. Driver 2 scatter plots, including the least squares line and r, of metric values and aspect scores 4, sorted from highest to lowest absolute r-value.

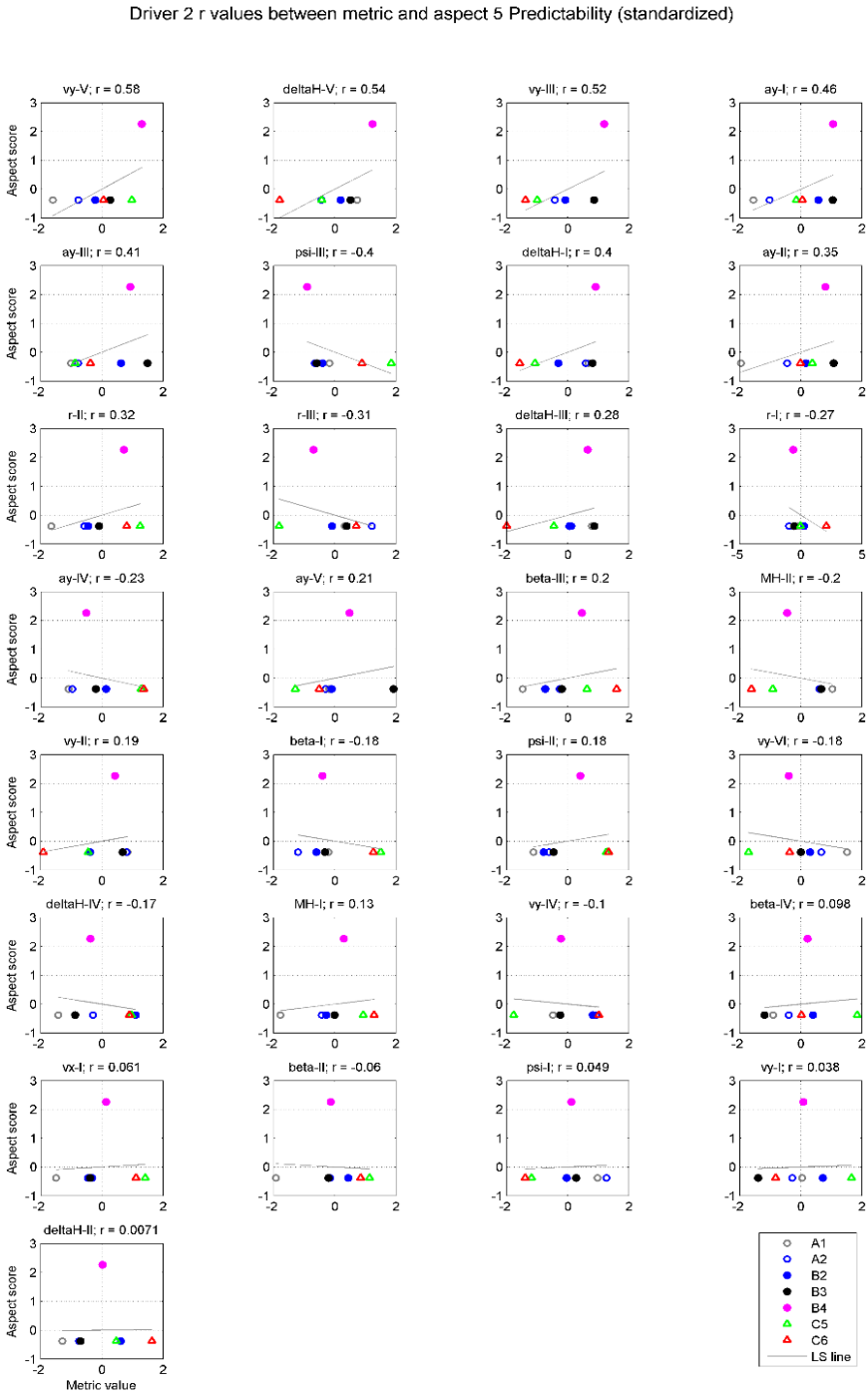


Figure 155. Driver 2 scatter plots, including the least squares line and r, of metric values and aspect scores 5, sorted from highest to lowest absolute r-value.

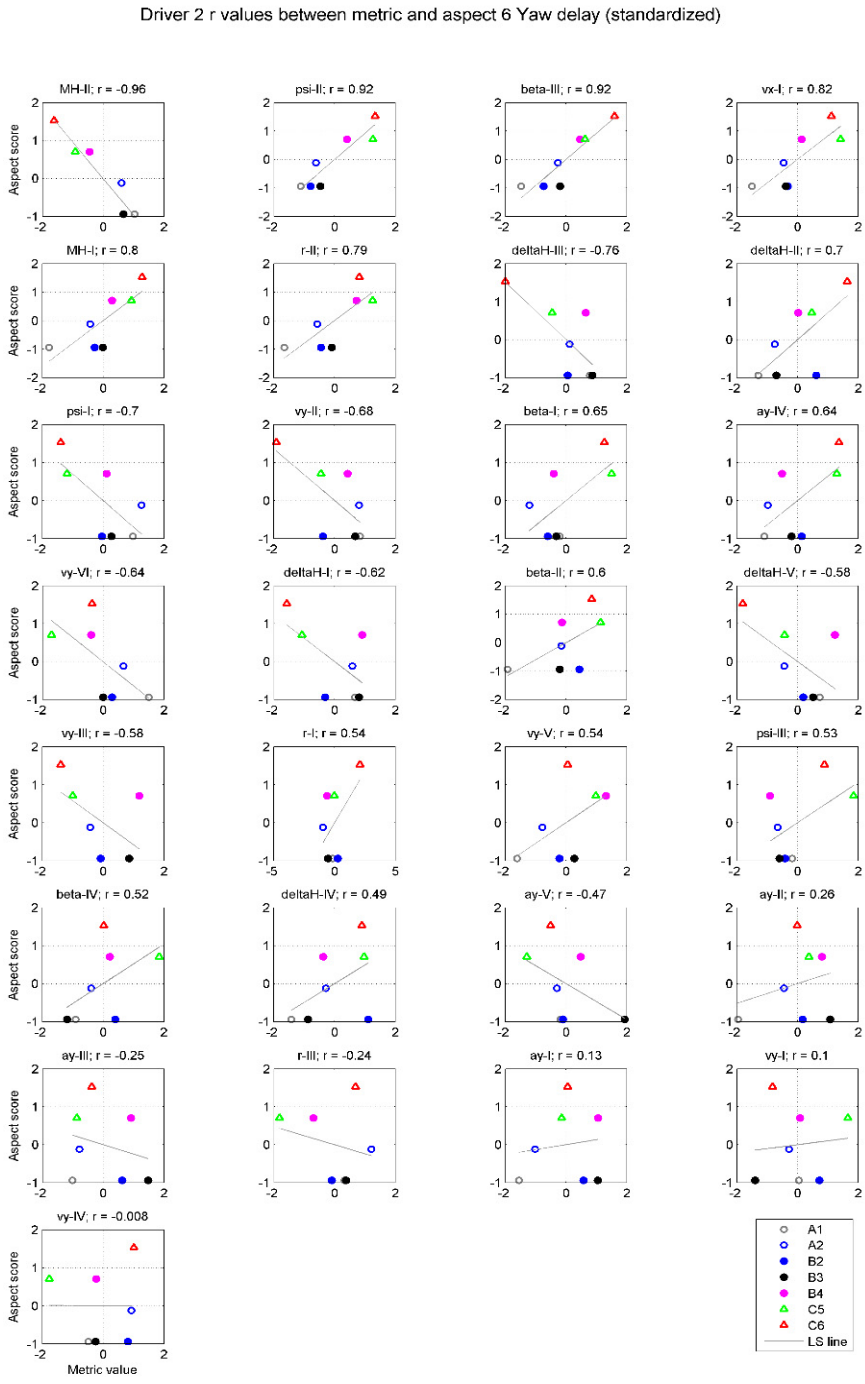


Figure 156. Driver 2 scatter plots, including the least squares line and r, of metric values and aspect scores 6, sorted from highest to lowest absolute r-value.

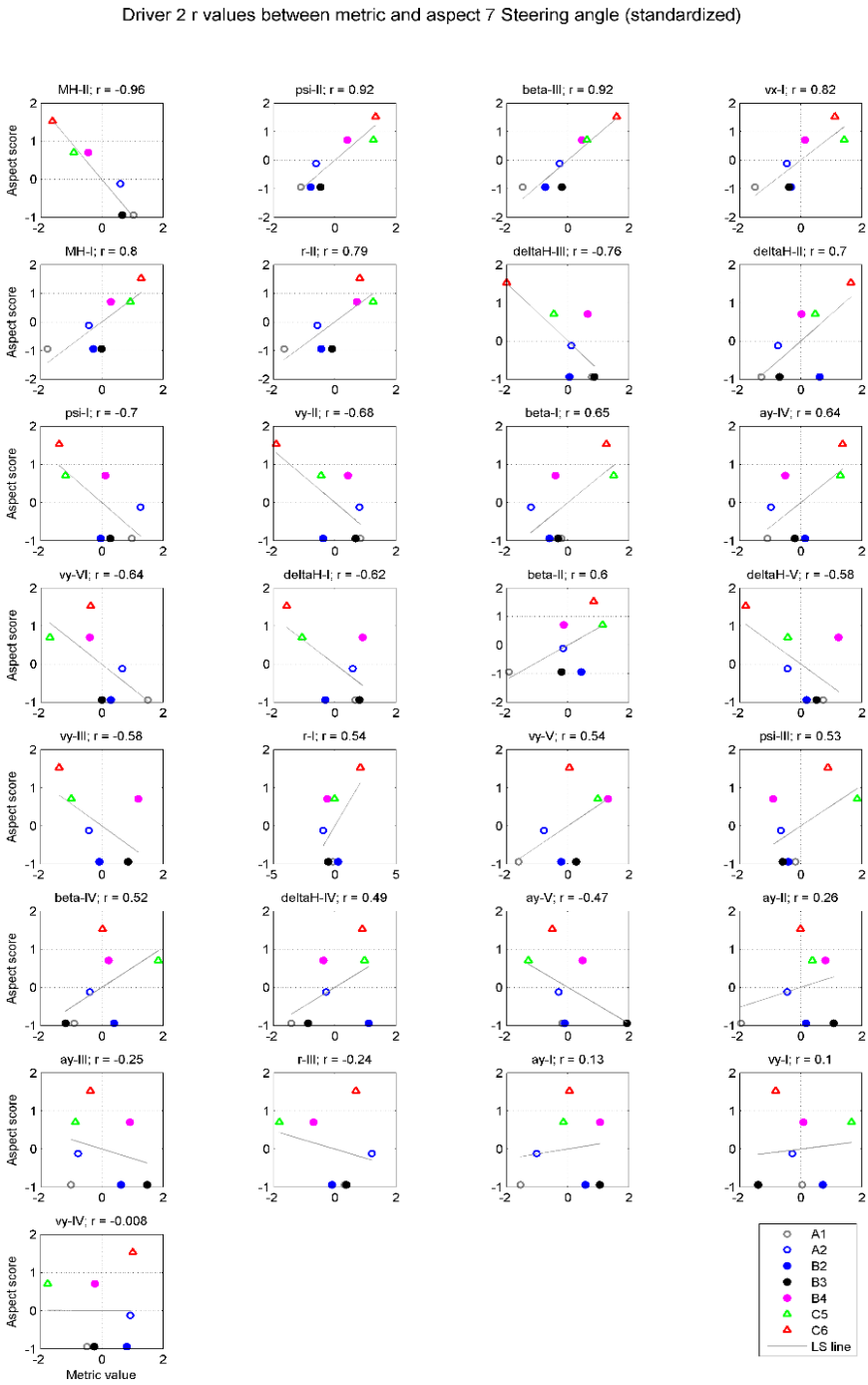


Figure 157. Driver 2 scatter plots, including the least squares line and r, of metric values and aspect scores 7, sorted from highest to lowest absolute r-value.

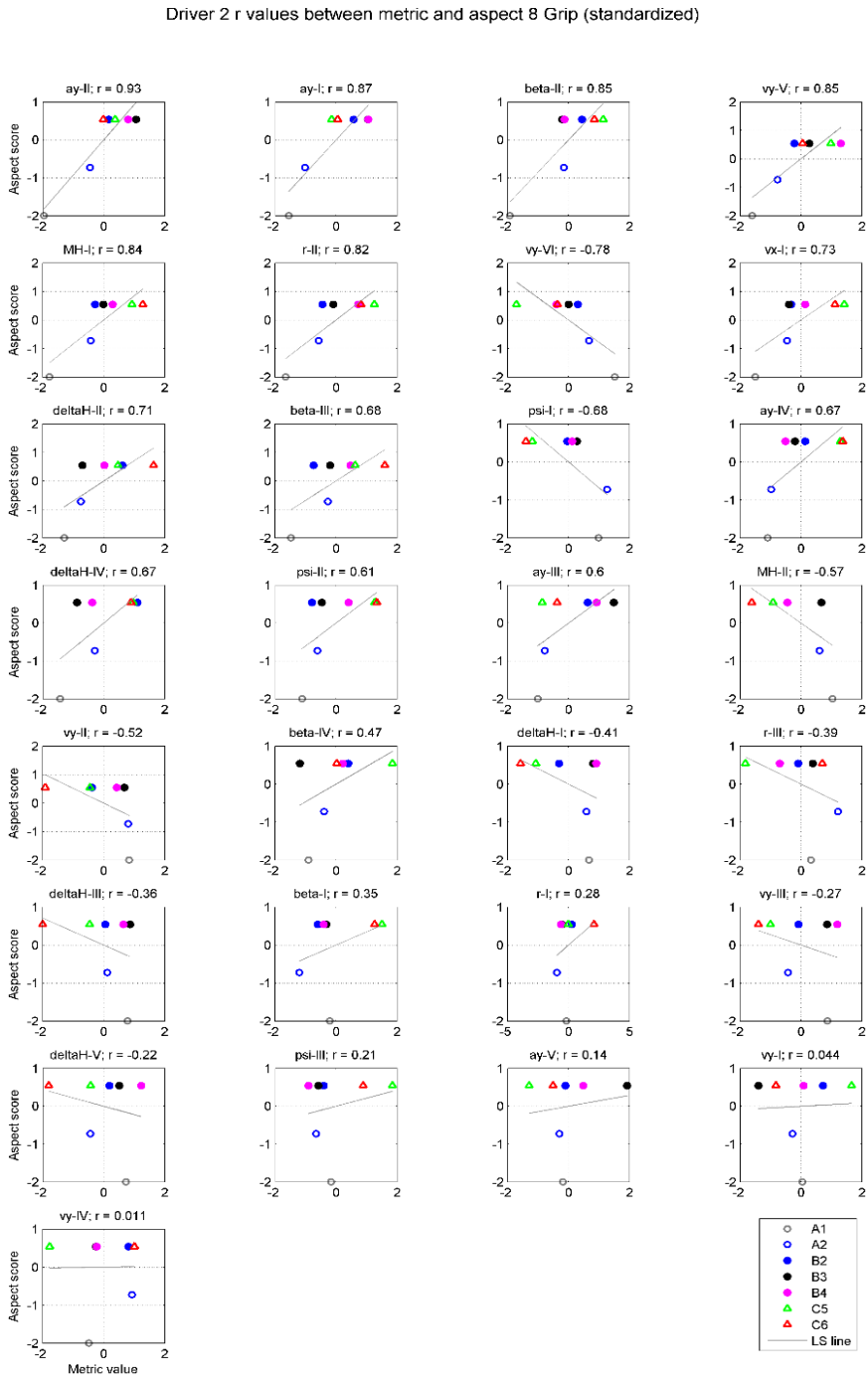


Figure 158. Driver 2 scatter plots, including the least squares line and r, of metric values and aspect scores 8, sorted from highest to lowest absolute r-value.

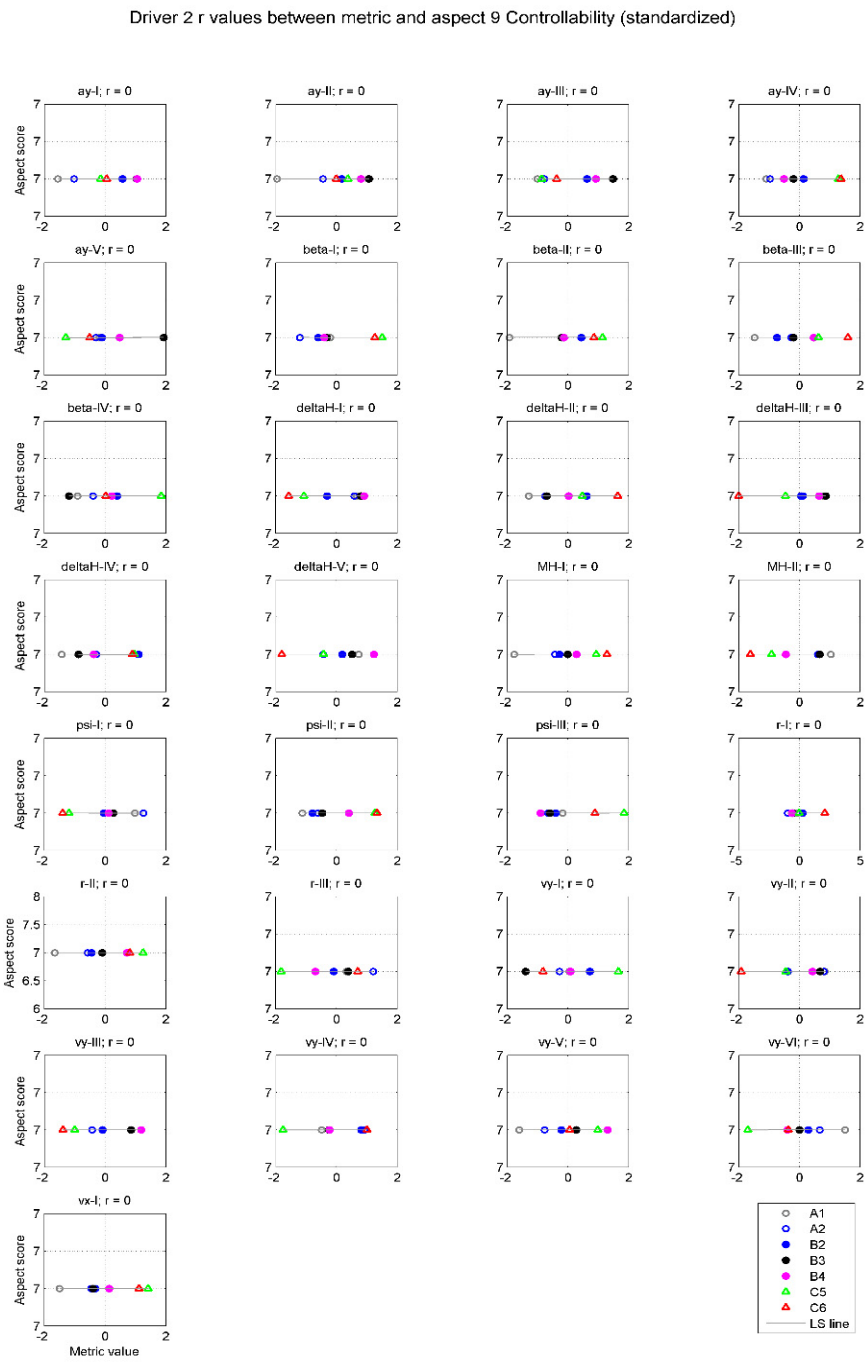


Figure 159. Driver 2 scatter plots, including the least squares line and r, of metric values and aspect scores 9, sorted from highest to lowest absolute r-value.

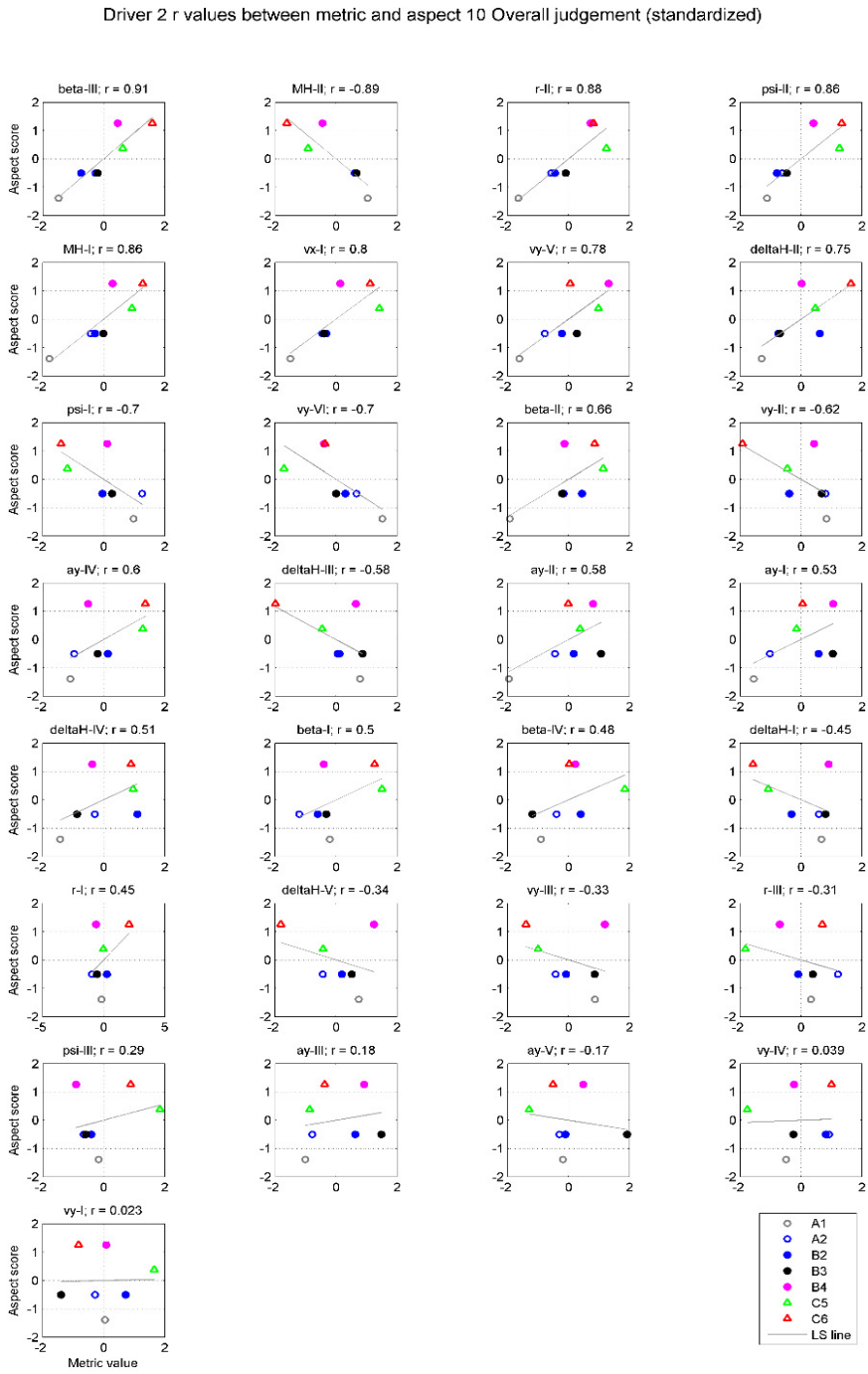


Figure 160. Driver 2 scatter plots, including the least squares line and r, of metric values and aspect scores 10, sorted from highest to lowest absolute r-value.

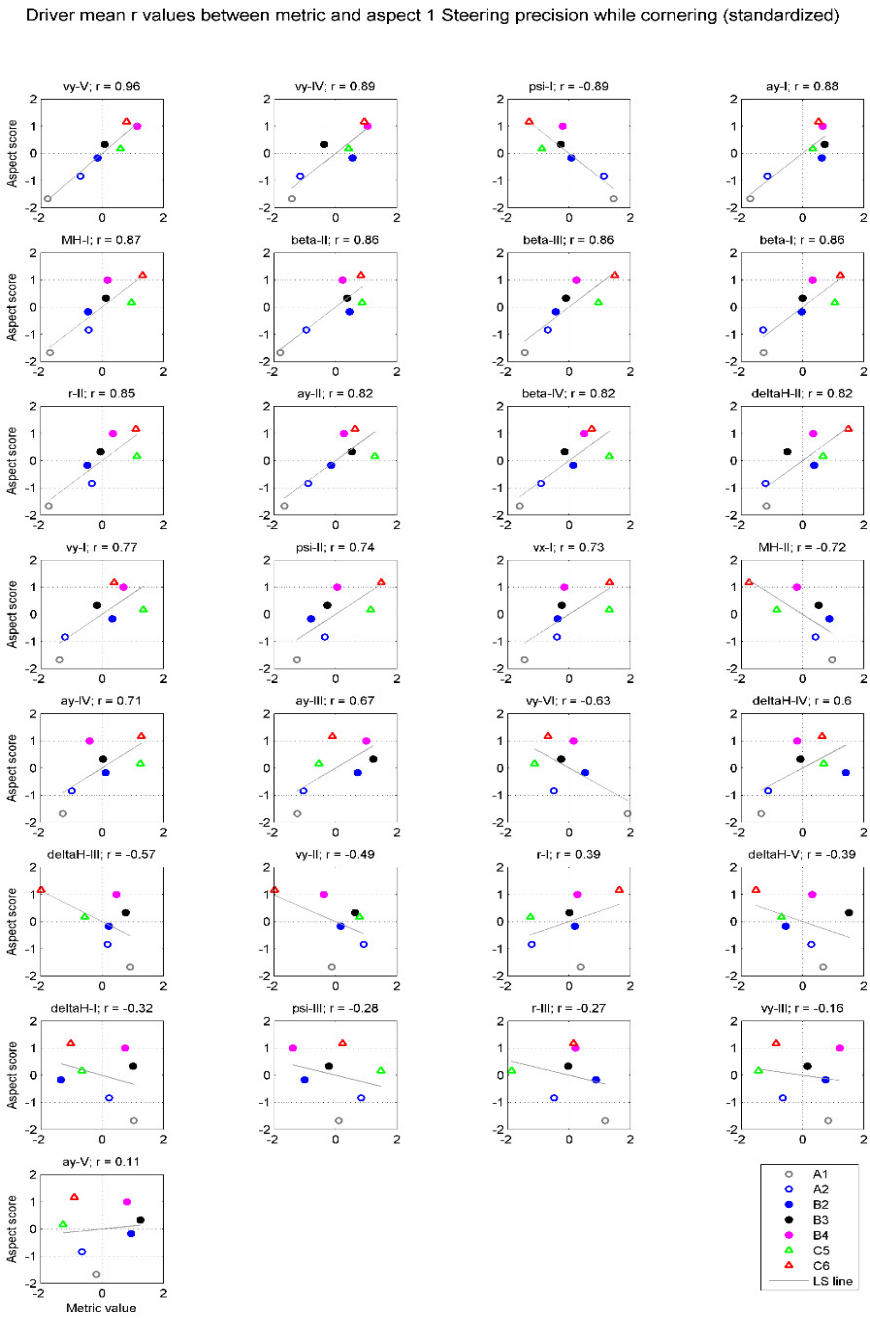


Figure 161. Driver mean scatter plots, including the least squares line and r, of metric values and aspect scores 1, sorted from highest to lowest absolute r-value.

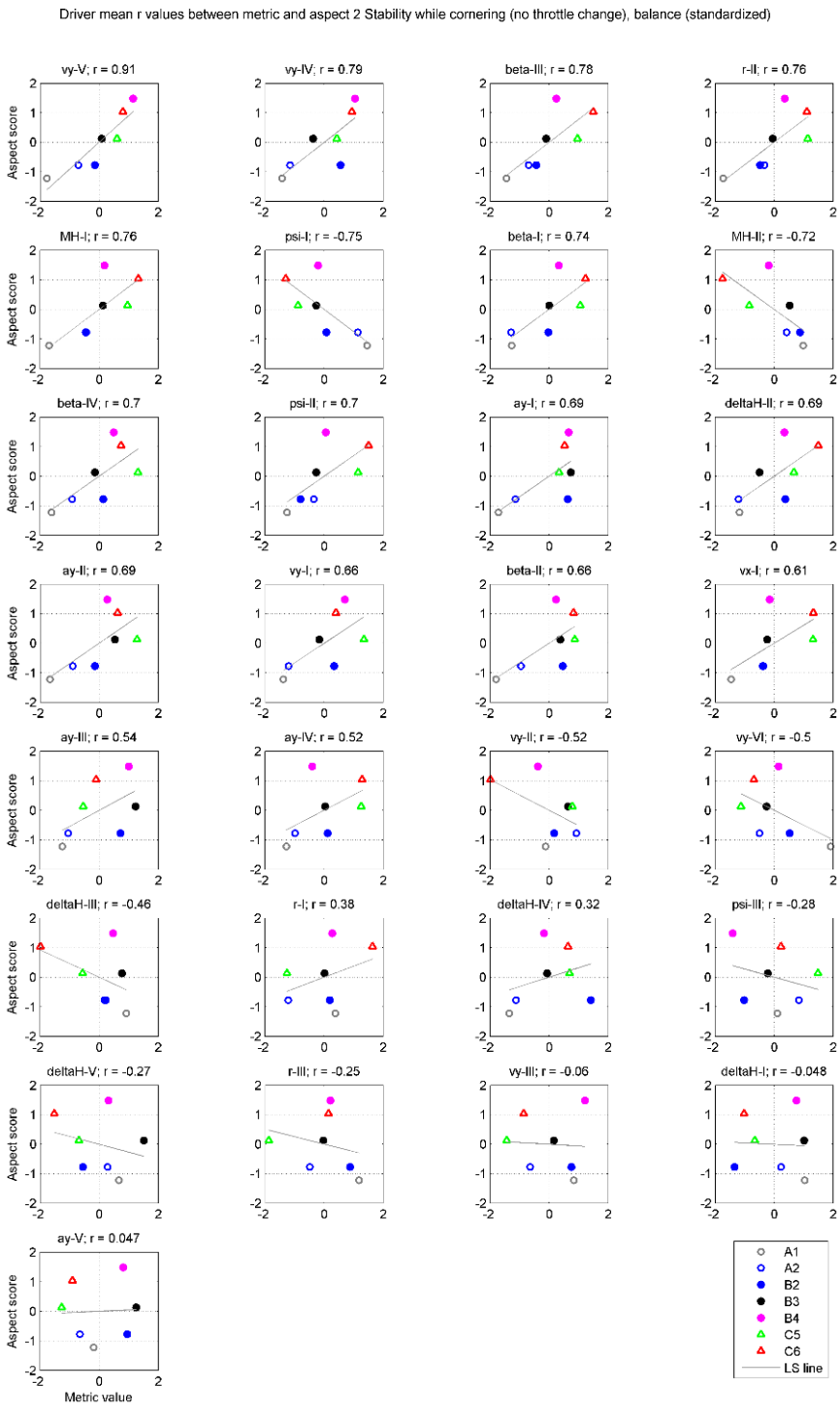


Figure 162. Driver mean scatter plots, including the least squares line and r, of metric values and aspect scores 2, sorted from highest to lowest absolute r-value.

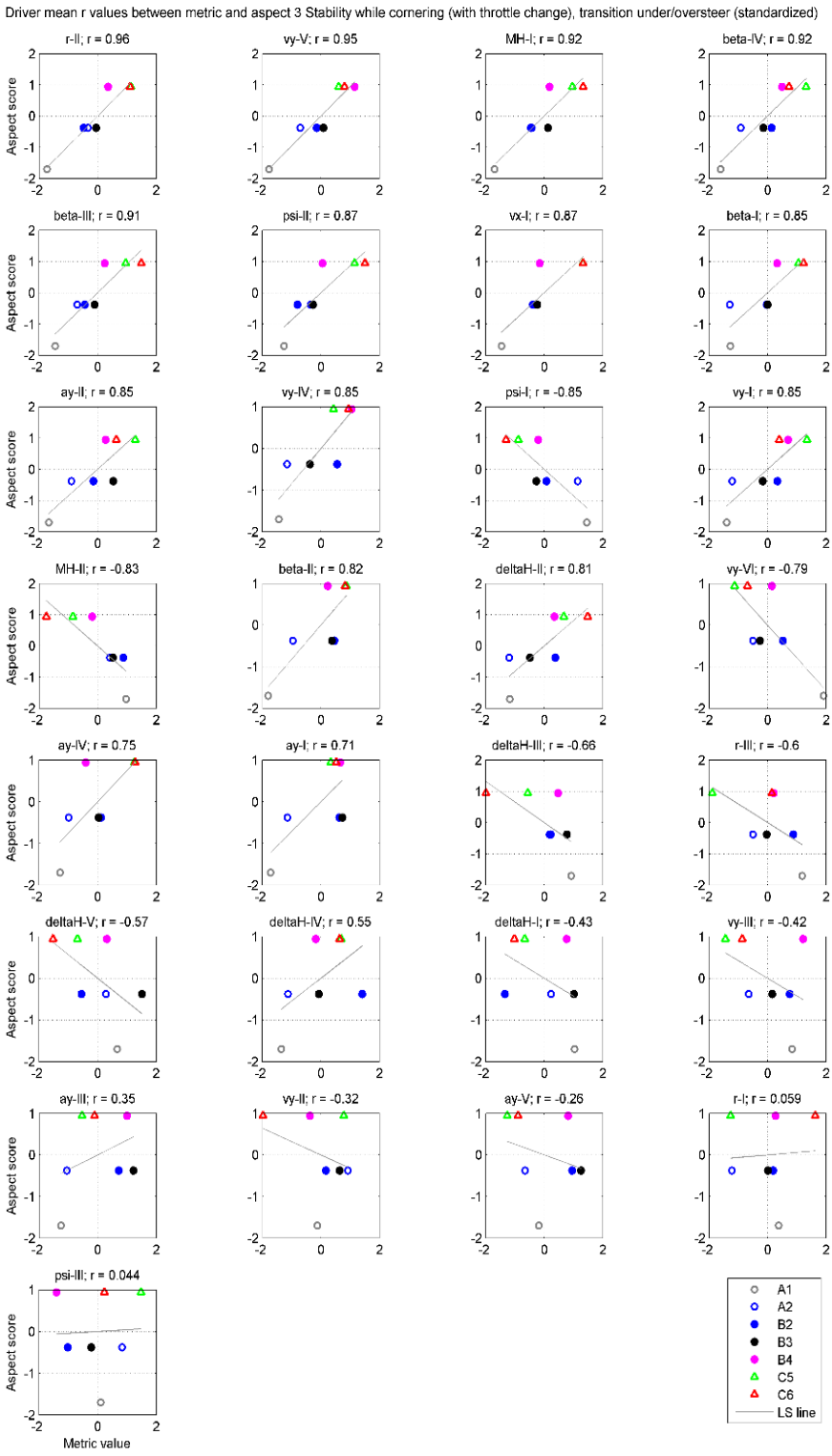


Figure 163. Driver mean scatter plots, including the least squares line and r, of metric values and aspect scores 3, sorted from highest to lowest absolute r-value.

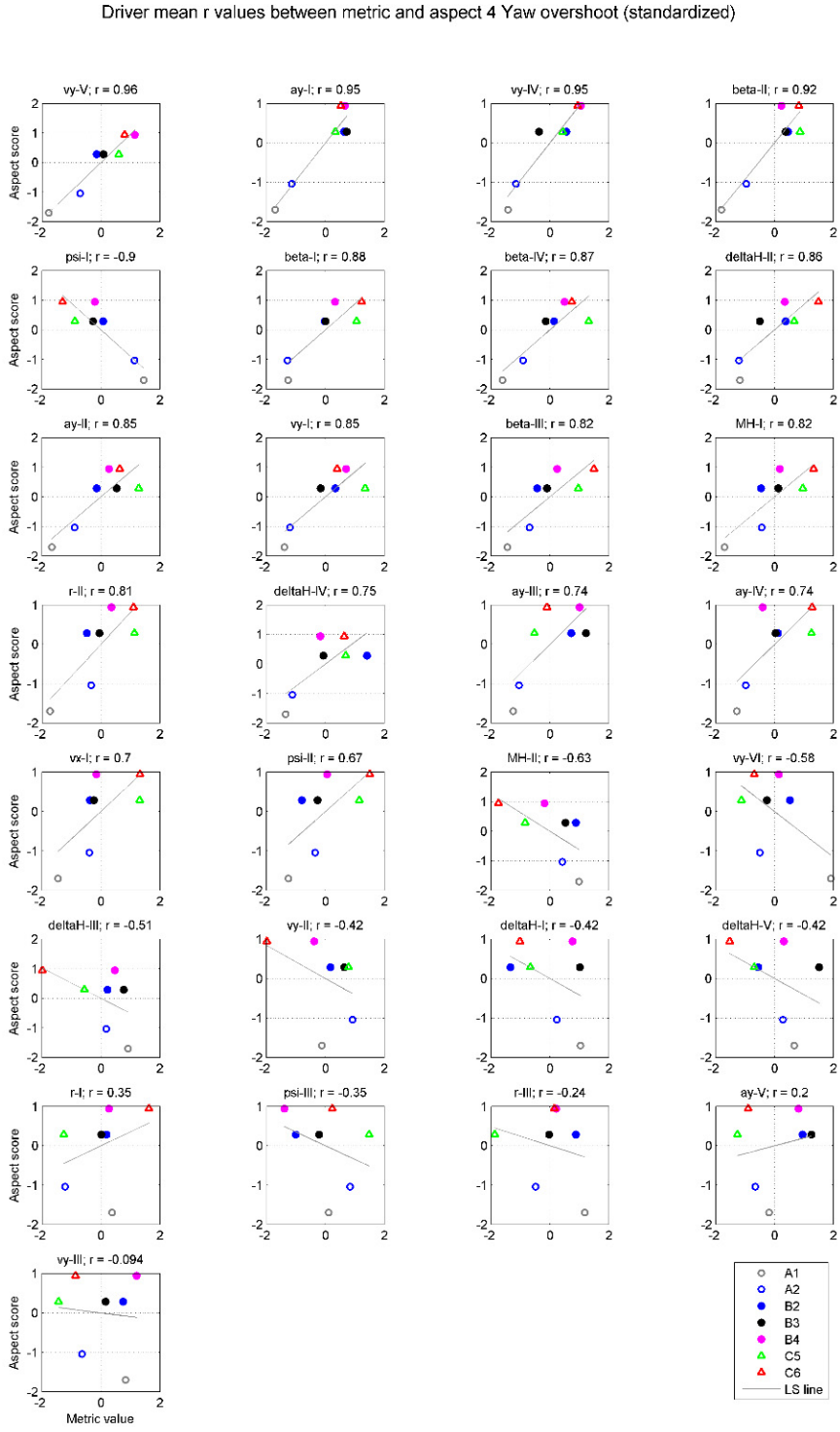


Figure 164. Driver mean scatter plots, including the least squares line and r, of metric values and aspect scores 4, sorted from highest to lowest absolute r-value.

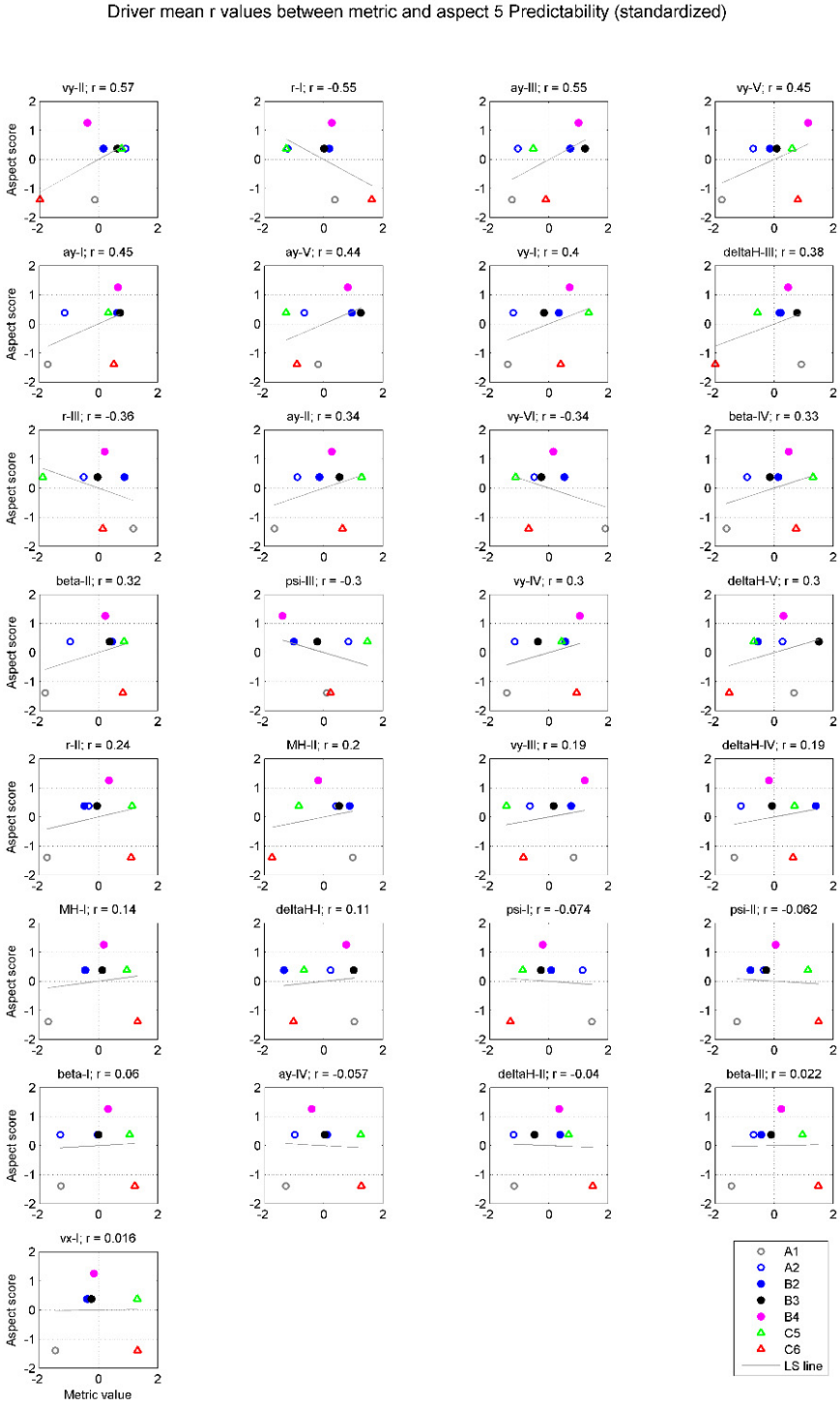


Figure 165. Driver mean scatter plots, including the least squares line and r, of metric values and aspect scores 5, sorted from highest to lowest absolute *r*-value.

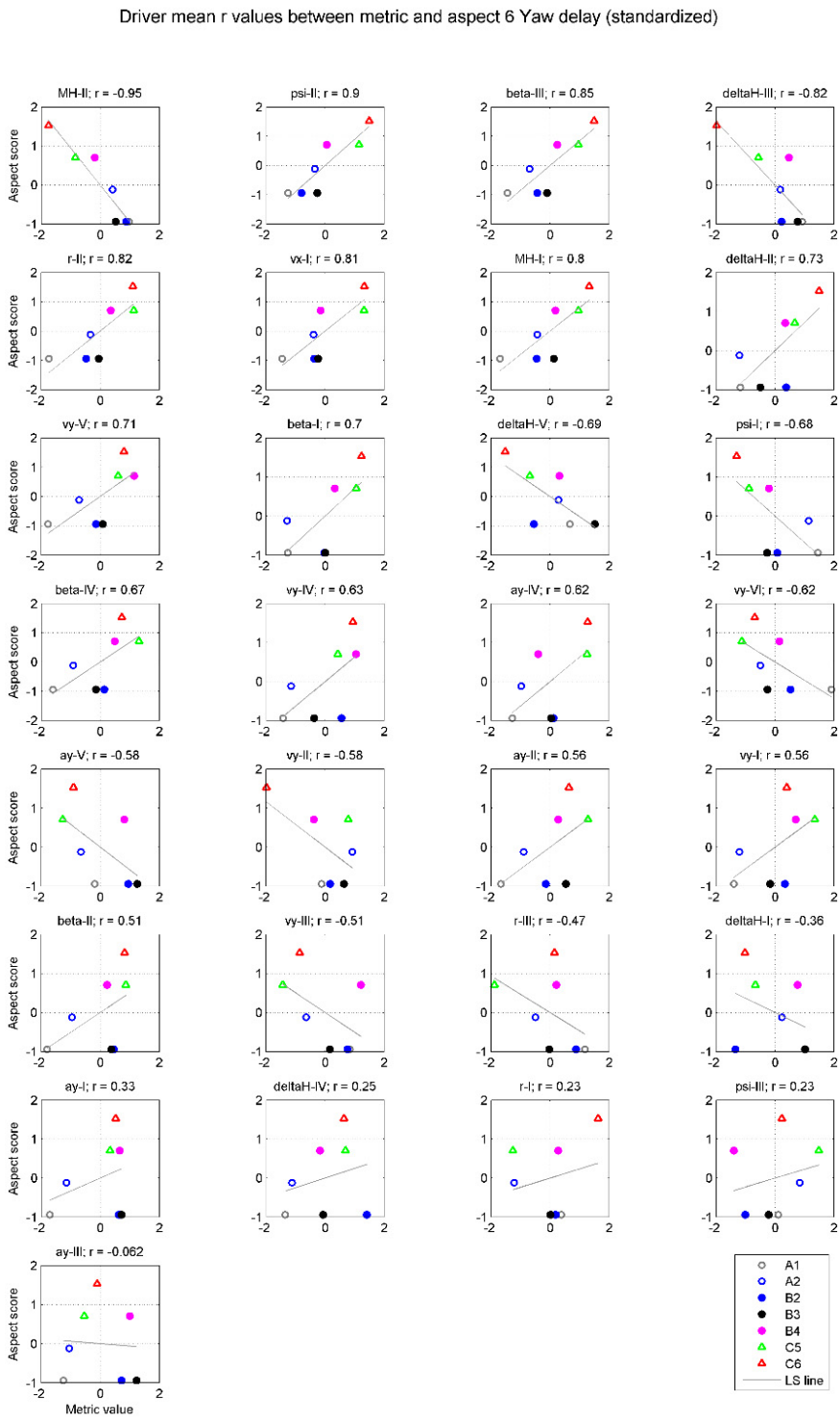


Figure 166. Driver mean scatter plots, including the least squares line and r, of metric values and aspect scores 6, sorted from highest to lowest absolute r-value.

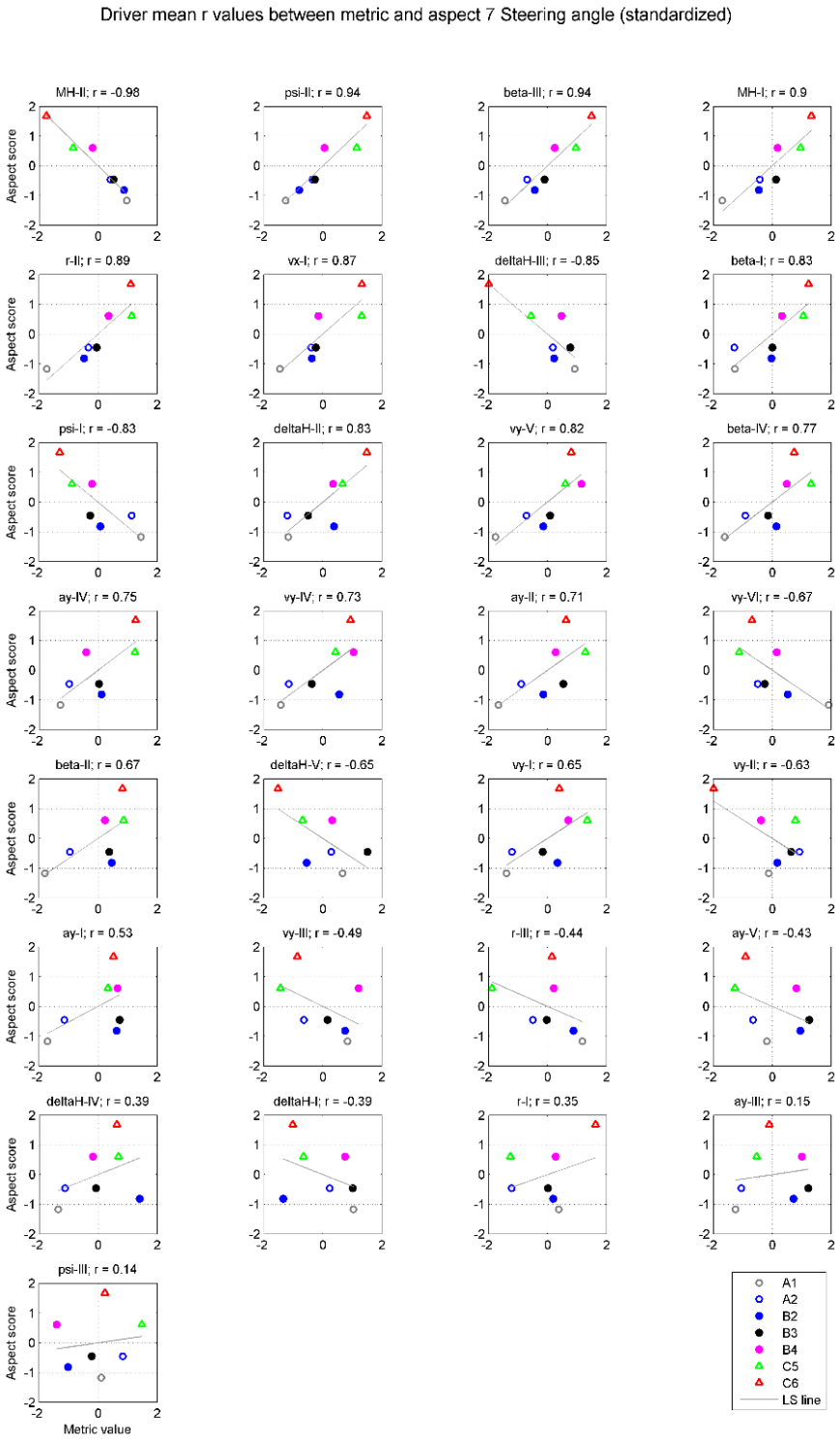


Figure 167. Driver mean scatter plots, including the least squares line and r, of metric values and aspect scores 7, sorted from highest to lowest absolute *r*-value.

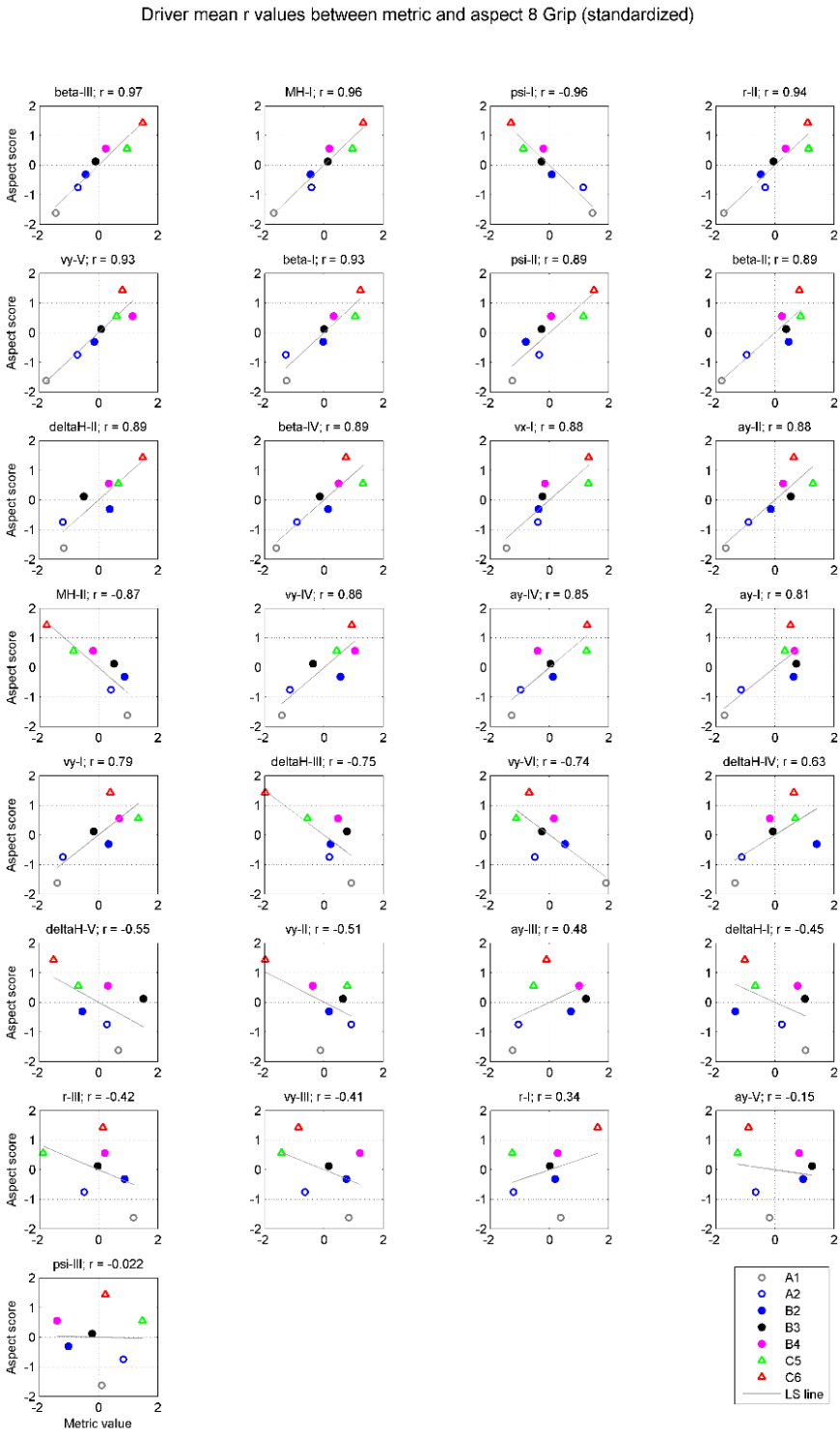


Figure 168. Driver mean scatter plots, including the least squares line and r, of metric values and aspect scores 8, sorted from highest to lowest absolute r-value.

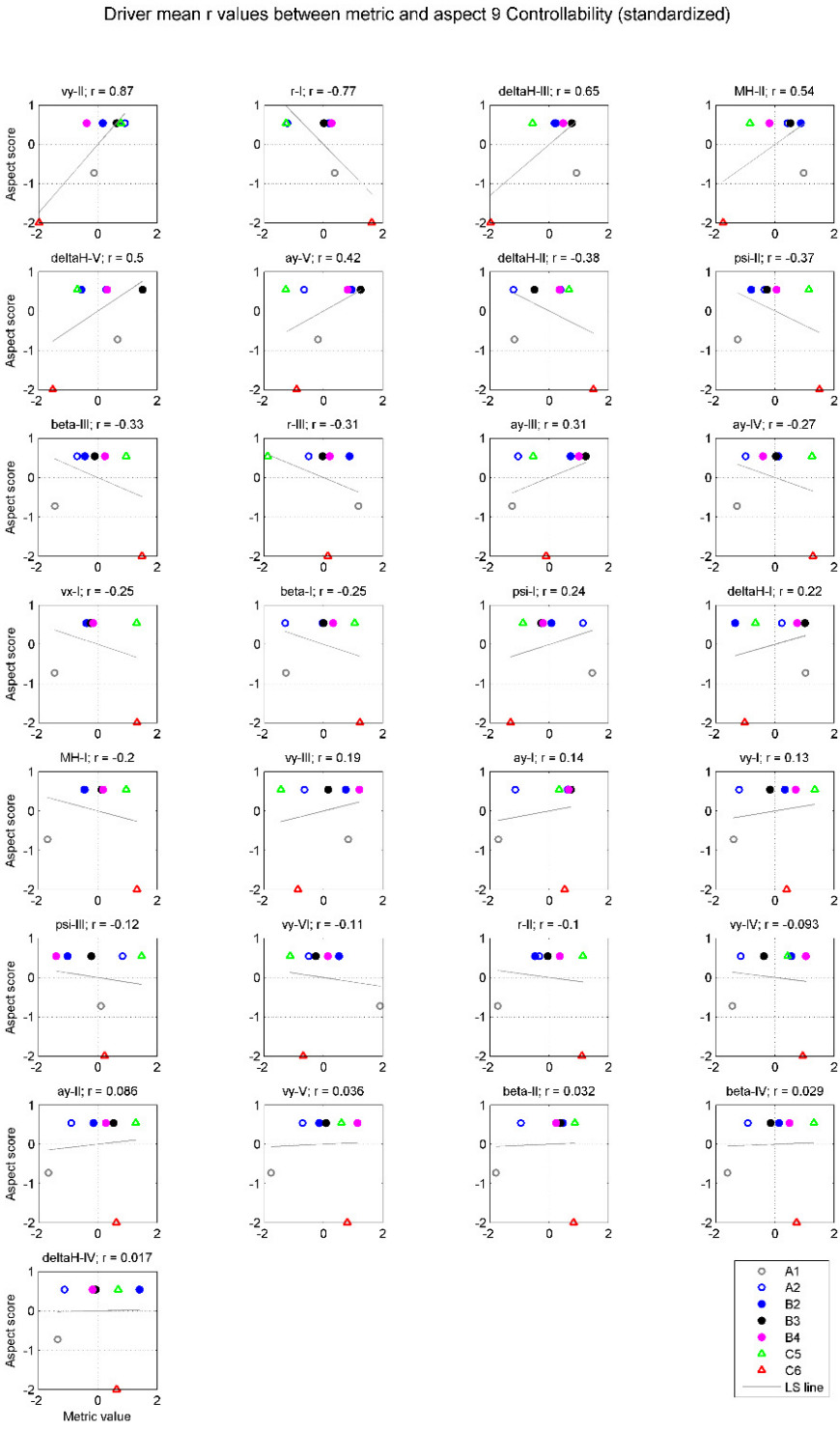


Figure 169. Driver mean scatter plots, including the least squares line and r, of metric values and aspect scores 9, sorted from highest to lowest absolute *r*-value.

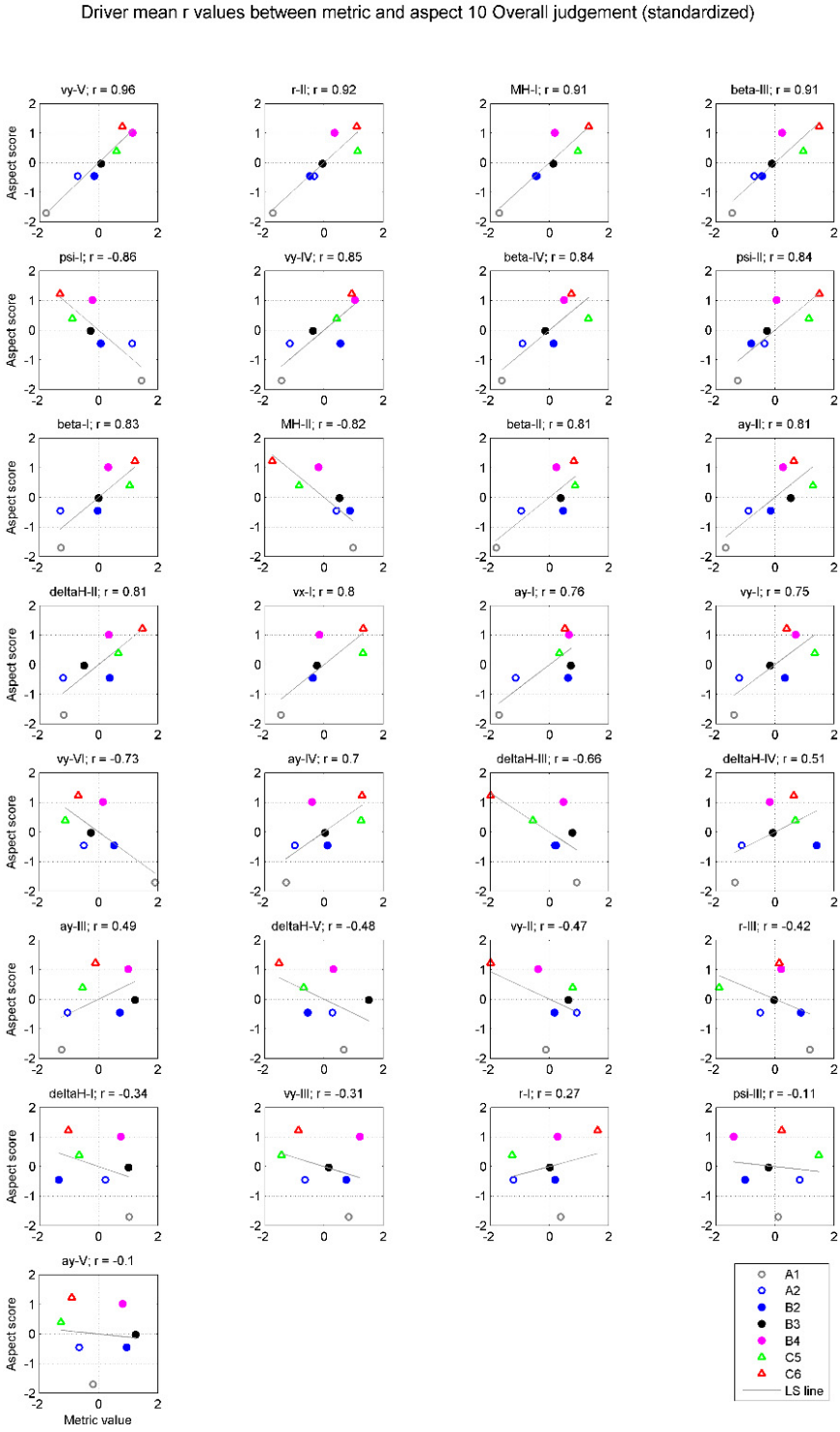


Figure 170. Driver mean scatter plots, including the least squares line and r, of metric values and aspect scores 10, sorted from highest to lowest absolute r-value.

Appendix 7 Mean Model Prediction

In this appendix the mean model prediction for every left out test assessment target scores are given as described in section 6.4.3, are given. Each assessment is identified with a letter-number combination. The letter indicates the batch the tire is driven in: A, B or C; the number indicates the tire: 1 to 6. In Tables 29 to 31 the mean model prediction scores are given for respectively driver 1, driver 2 and driver mean.

Table 29. Driver 1 mean model prediction values and RMSE for every test assessment.

Test assessment / Aspect	A1	A2	B2	B3	B4	C5	C6
1. Steering precision while cornering	6.83	6.67	6.67	6.54	6.46	6.50	6.33
2. Stability while cornering (no throttle change)	7.00	6.92	6.92	6.75	6.58	6.75	6.58
3. Stability while cornering (throttle change)	7.33	7.33	7.33	7.33	7.33	7.17	7.17
4. Yaw overshoot	7.83	7.83	7.67	7.67	7.67	7.67	7.67
5. Predictability	7.83	7.67	7.67	7.67	7.67	7.67	7.83
6. Yaw delay	7.00	7.00	7.00	7.00	7.00	7.00	7.00
7. Steering angle	6.75	6.67	6.67	6.58	6.50	6.50	6.33
8. Grip	6.75	6.67	6.67	6.58	6.50	6.50	6.33
9. Controllability	7.83	7.75	7.75	7.75	7.75	7.75	7.92
10. Overall judgement	7.46	7.29	7.29	7.21	7.17	7.21	7.13
RMSE (calculated over all actual aspect scores and mean model aspect scores)	**1.12**	**0.57**	**0.52**	**0.25**	**0.62**	**0.43**	**1.16**

Table 30. Driver 2 mean model prediction values and RMSE for every test assessment.

Aspect \ Test assessment	A1	A2	B2	B3	B4	C5	C6
1. Steering precision while cornering	6.83	6.79	6.63	6.63	6.54	6.71	6.63
2. Stability while cornering (no throttle change)	7.08	7.08	7.08	7.08	7.00	7.08	7.08
3. Stability while cornering (throttle change)	6.92	6.83	6.83	6.83	6.75	6.92	6.92
4. Yaw overshoot	6.67	6.58	6.58	6.58	6.50	6.58	6.50
5. Predictability	7.08	7.08	7.08	7.08	7.00	7.08	7.08
6. Yaw delay	6.67	6.58	6.67	6.67	6.50	6.50	6.42
7. Steering angle	6.67	6.58	6.67	6.67	6.50	6.50	6.42
8. Grip	6.92	6.83	6.75	6.75	6.75	6.75	6.75
9. Controllability	7.00	7.00	7.00	7.00	7.00	7.00	7.00
10. Overall judgement	6.92	6.83	6.83	6.83	6.67	6.75	6.67
RMSE (calculated over all actual aspect scores and mean model aspect scores)	**0.66**	**0.30**	**0.35**	**0.35**	**0.59**	**0.29**	**0.61**

Table 31. Driver mean mean model prediction values and RMSE for every test assessment.

Aspect \ Test assessment	A1	A2	B2	B3	B4	C5	C6
1. Steering precision while cornering	6.83	6.73	6.65	6.58	6.50	6.60	6.48
2. Stability while cornering (no throttle change)	7.04	7.00	7.00	6.92	6.79	6.92	6.83
3. Stability while cornering (throttle change)	7.13	7.08	7.08	7.08	7.04	7.04	7.04
4. Yaw overshoot	7.25	7.21	7.13	7.13	7.08	7.13	7.08
5. Predictability	7.46	7.38	7.38	7.38	7.33	7.38	7.46
6. Yaw delay	6.83	6.79	6.83	6.83	6.75	6.75	6.71
7. Steering angle	6.71	6.63	6.67	6.63	6.50	6.50	6.38
8. Grip	6.83	6.75	6.71	6.67	6.63	6.63	6.54
9. Controllability	7.42	7.38	7.38	7.38	7.38	7.38	7.46
10. Overall judgement	7.19	7.06	7.06	7.02	6.92	6.98	6.90
RMSE (calculated over all actual aspect scores and mean model aspect scores)	**0.85**	**0.39**	**0.32**	**0.20**	**0.55**	**0.25**	**0.77**

Appendix 8 Prediction Performance Plots GRNN Test Assessments

In this appendix, GRNN prediction performance plots are presented, showing the RMSE between the actual and predicted output for the test input of the GRNN as function of the GRNN parameters number of predictors used and spread value. Fig. 171 to 173 show for each driver the surface plots for every test input assessments, which is left out in the learning phase of the GRNN. The same data is shown in contour plots in Fig. 174 to 176.

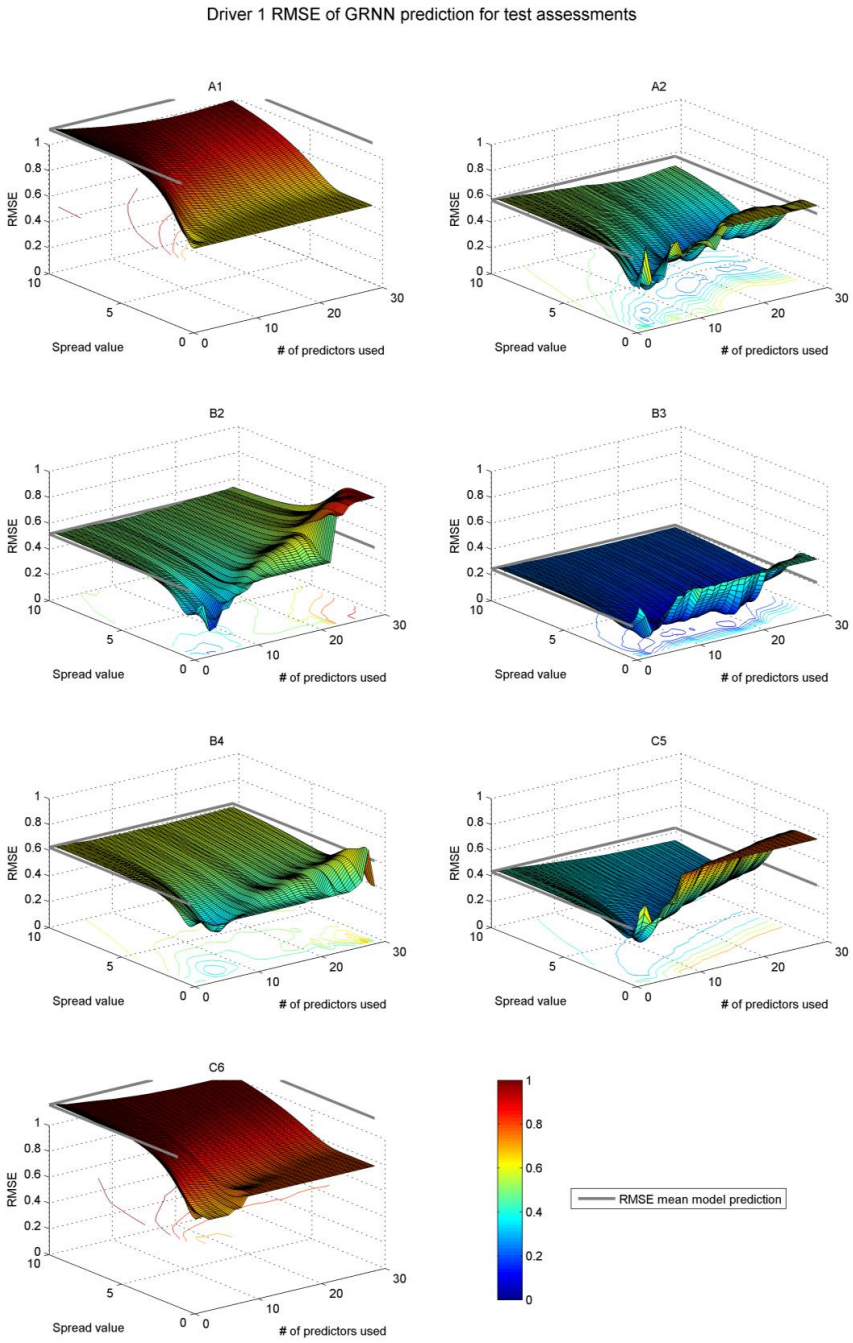


Figure 171. Driver 1 surface plots of the GRNN RMSE for all test input assessment as function of GRNN parameters spread value and number of predictors used.

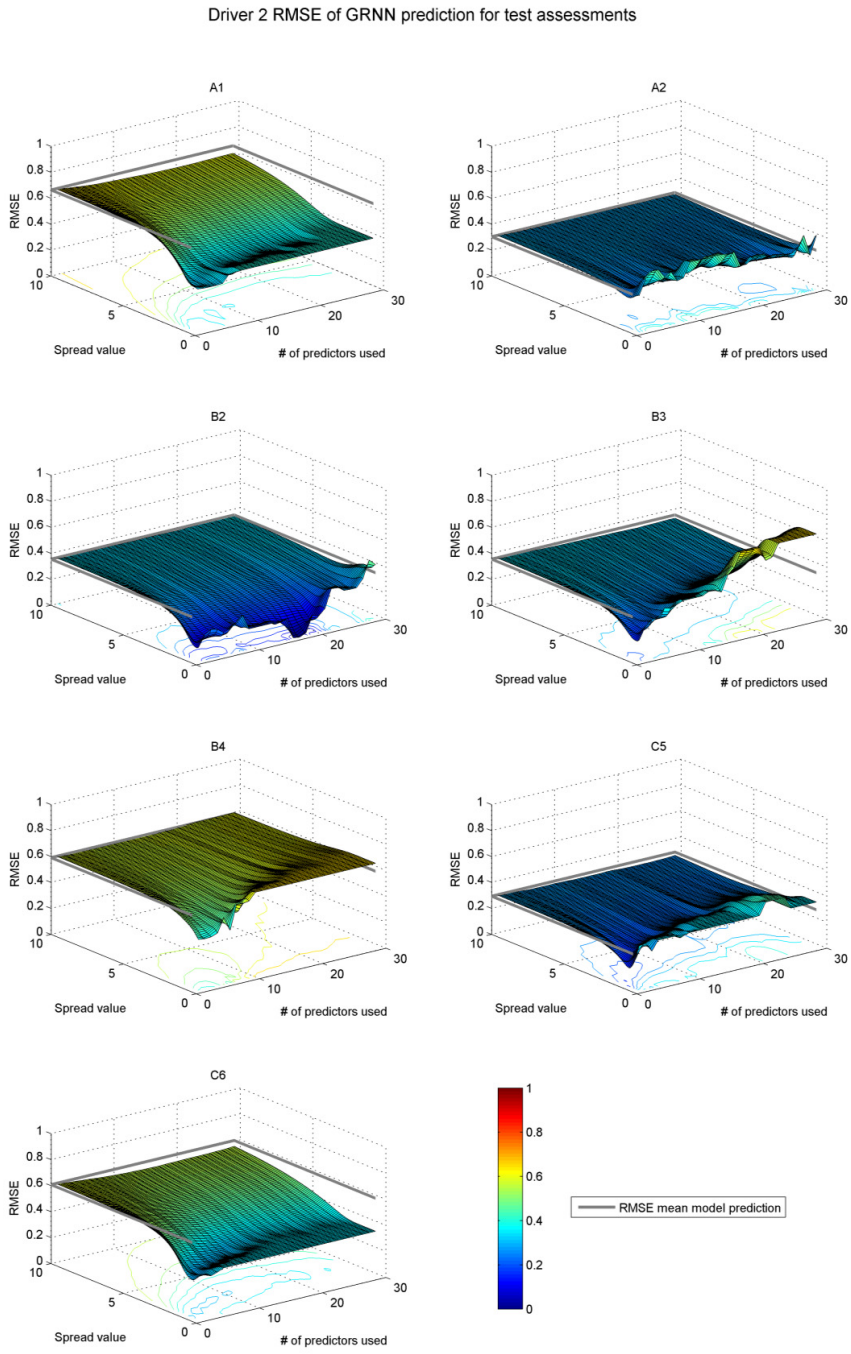


Figure 172. Driver 2 surface plots of the GRNN RMSE for all test input assessment as function of GRNN parameters spread value and number of predictors used.

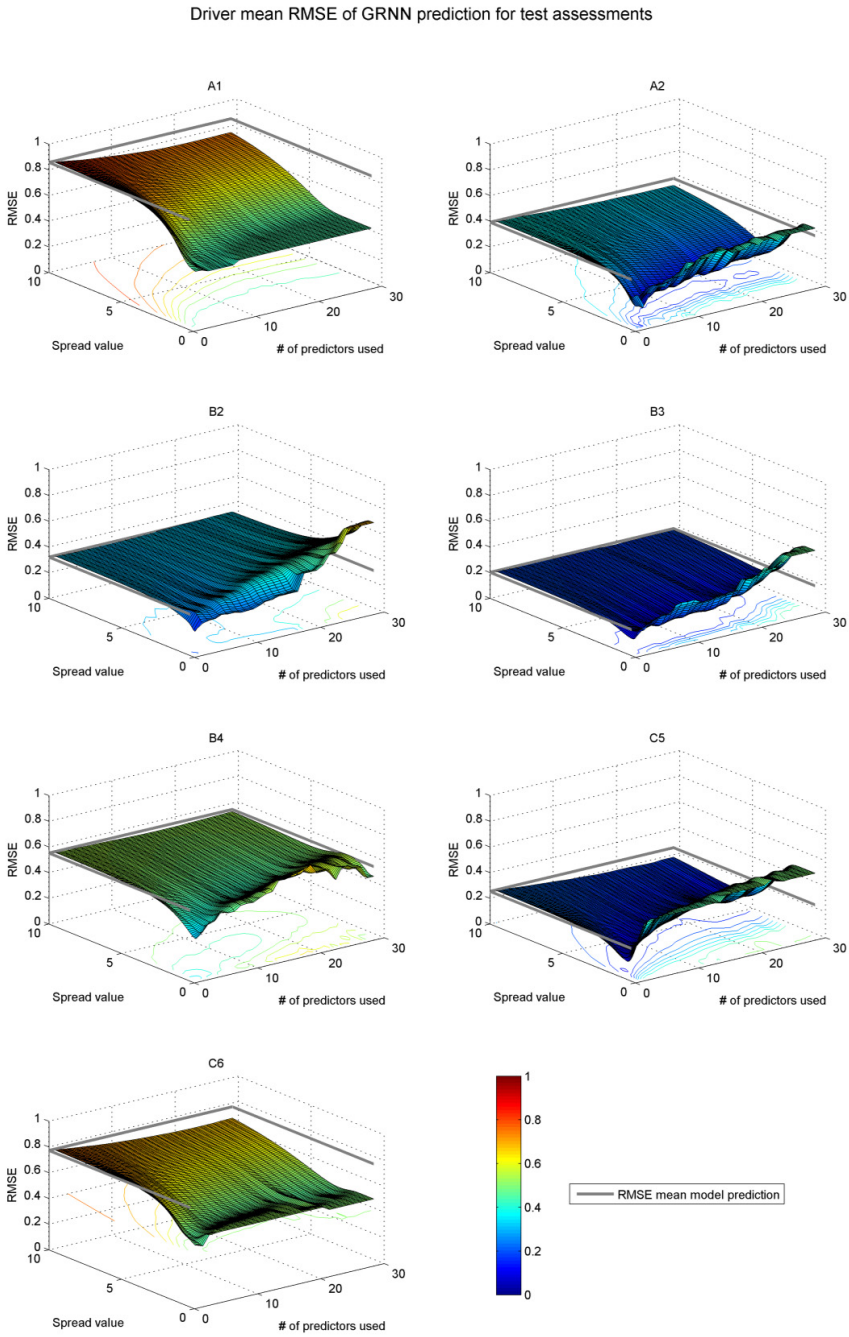


Figure 173. Driver mean surface plots of the GRNN RMSE for all test input assessment as function of GRNN parameters spread value and number of predictors used.

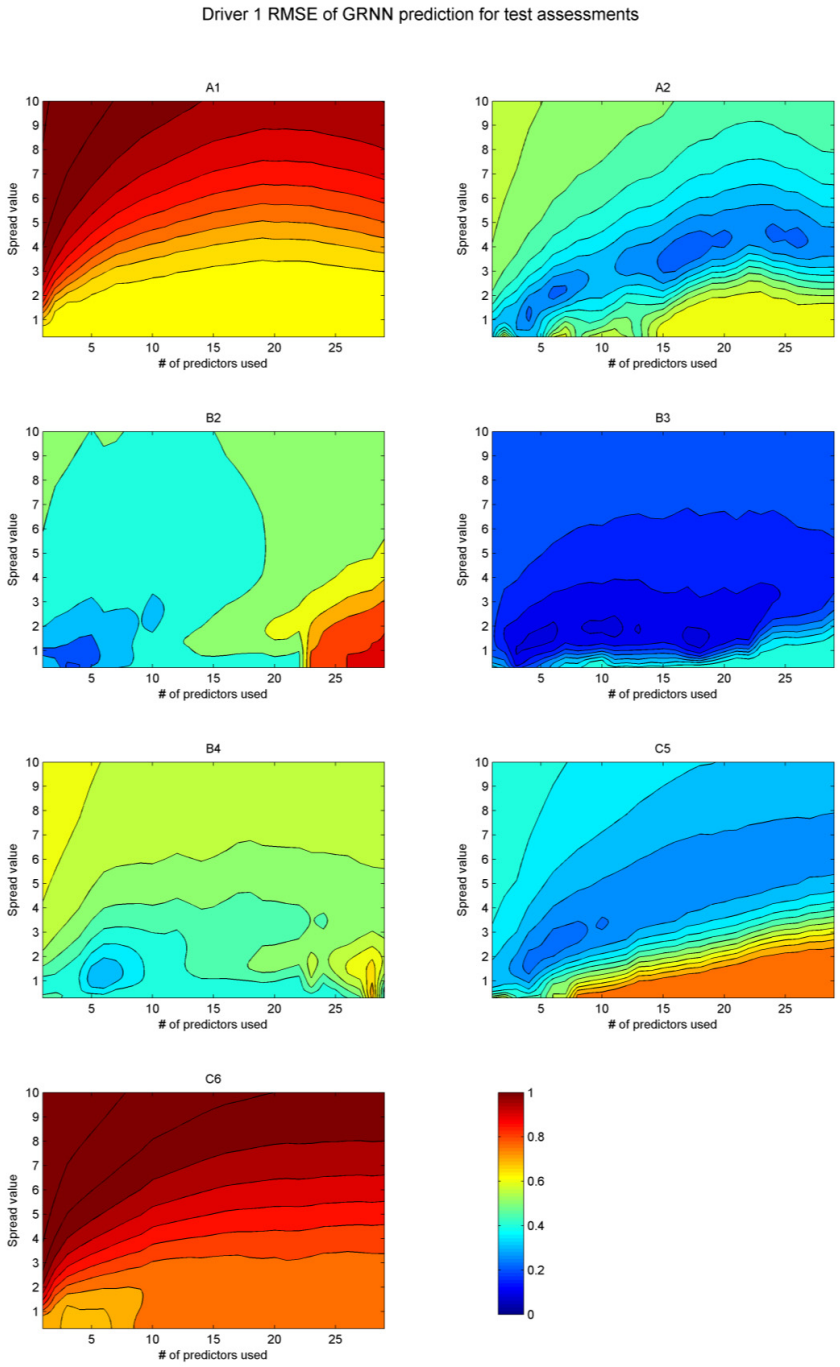


Figure 174. Driver 1 contour plots of the GRNN RMSE for all test input assessment as function of GRNN parameters spread value and number of predictors used.

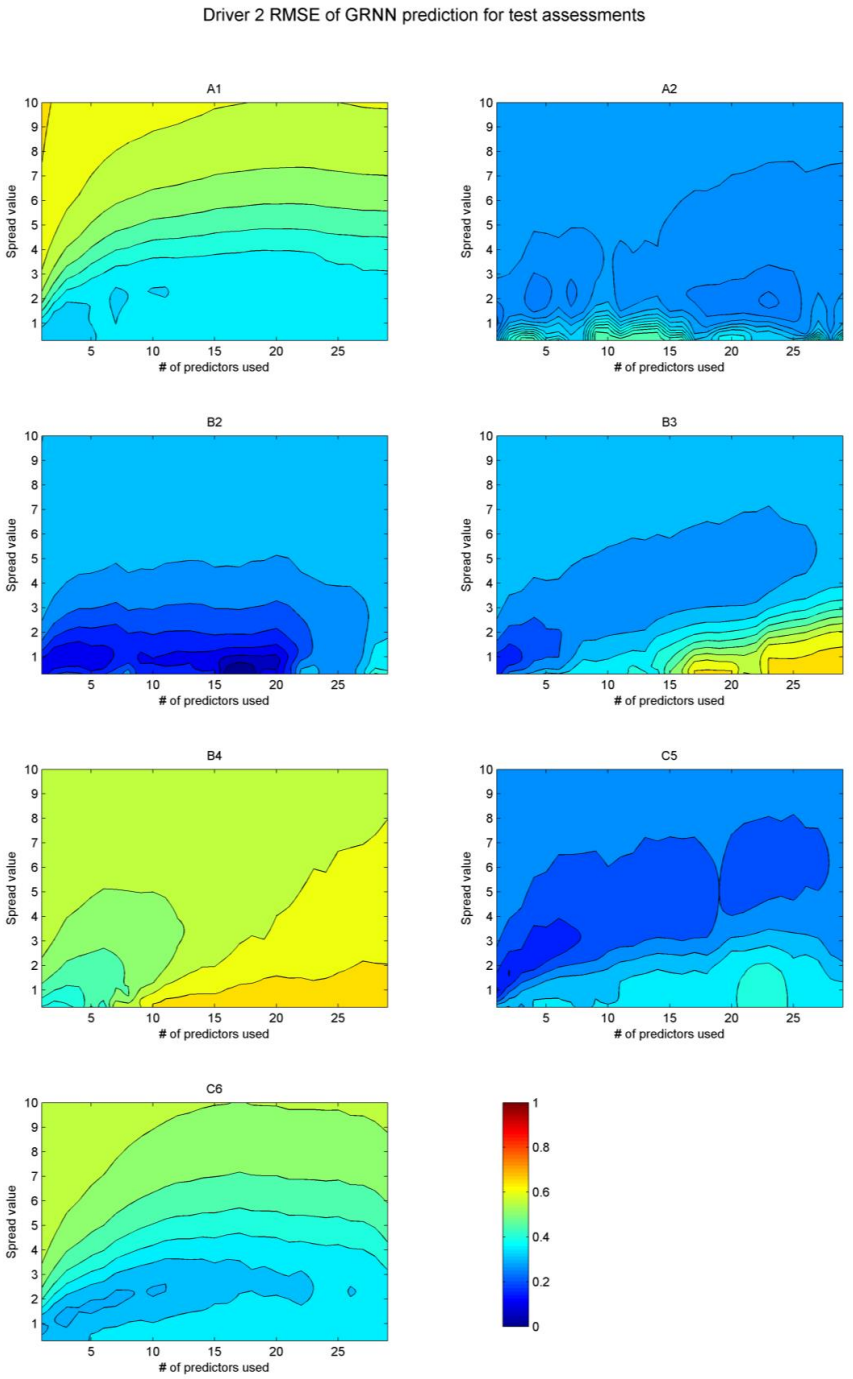


Figure 175. Driver 2 contour plots of the GRNN RMSE for all test input assessment as function of GRNN parameters spread value and number of predictors used.

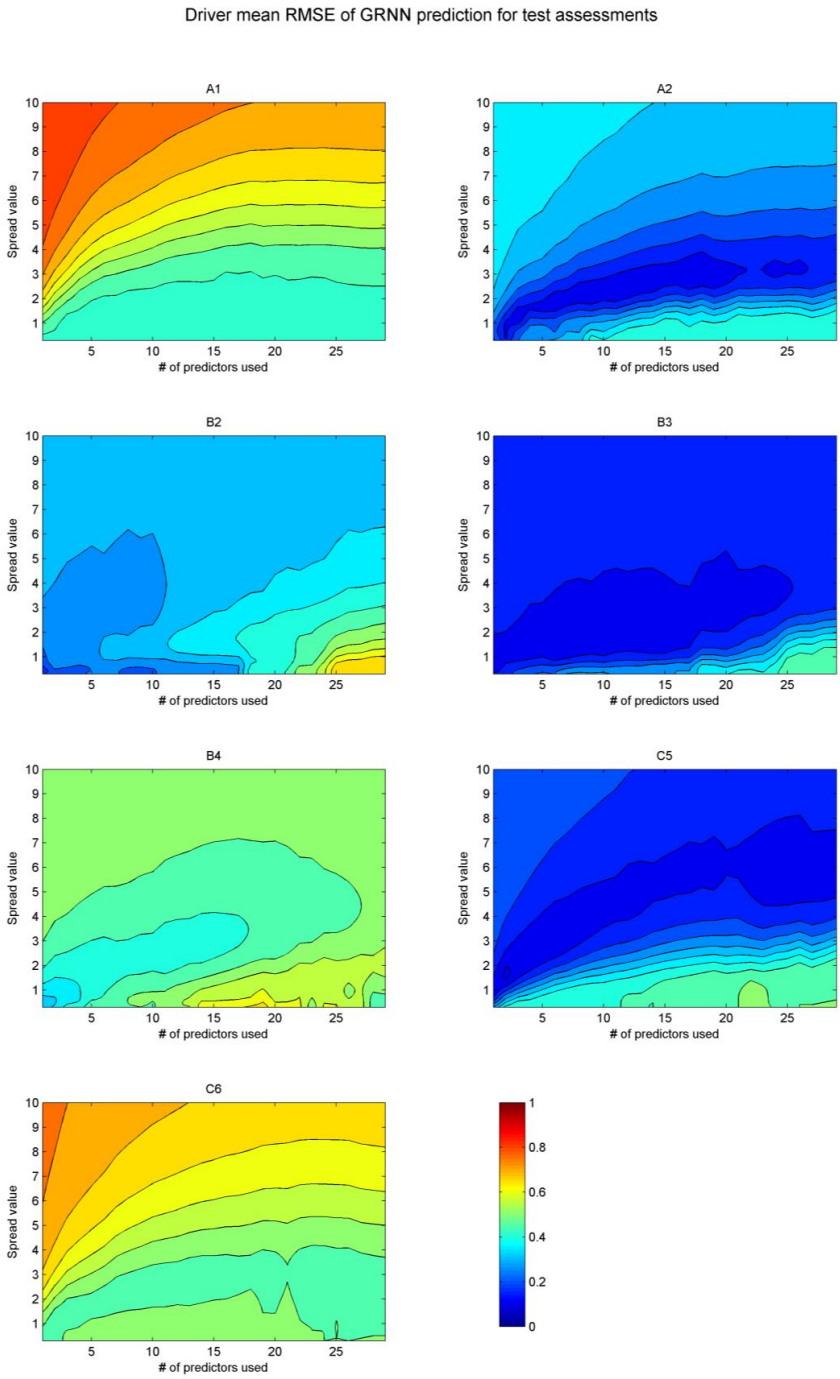


Figure 176. Driver mean contour plots of the GRNN RMSE for all test input assessment as function of GRNN parameters spread value and number of predictors used.

Appendix 9 Optimal Prediction Plots GRNN Test Assessments

In this appendix, the GRNN optimal prediction plots are presented, showing the GRNN predictions for the test input with the lowest RMSE, together with the corresponding number of predictors used and the spread value. Fig. 171 to 173 show for each driver the plots with actual, mean model predicted and GRNN predicted aspect scores for every test input assessments, which is left out in the learning phase of the GRNN.

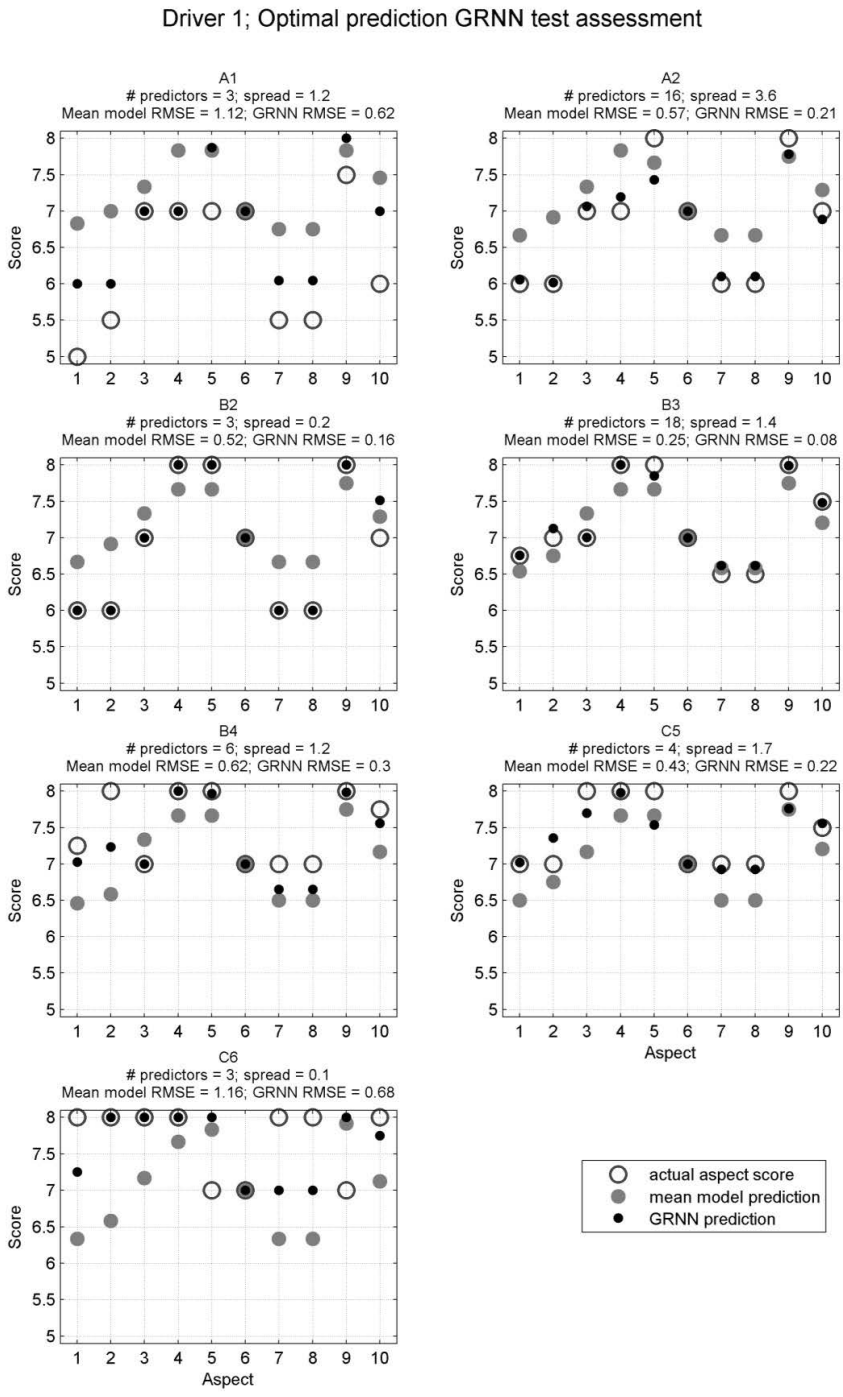


Figure 177. Driver 1 plots showing the actual, mean model and GRNN optimal aspect score for all test input assessment.

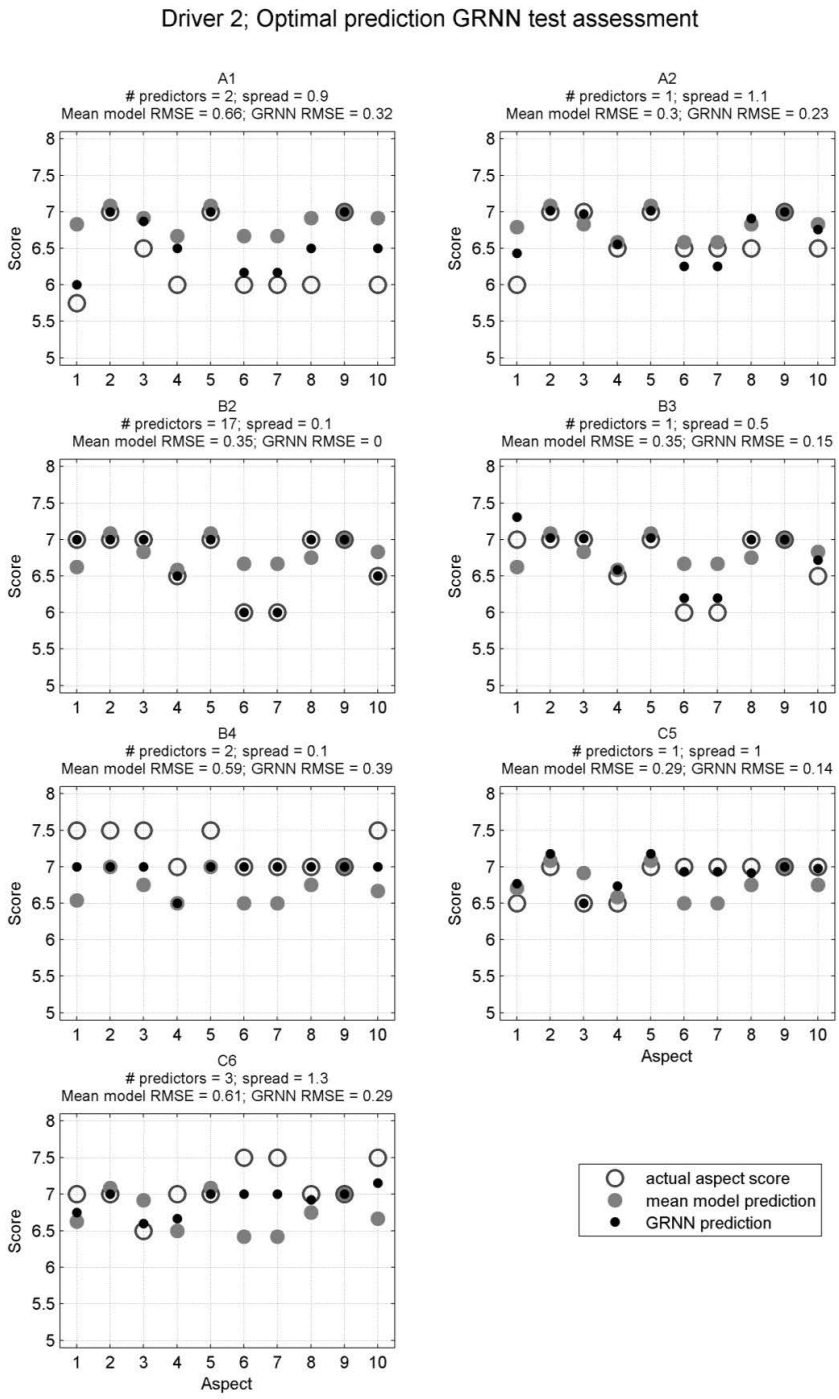


Figure 178. Driver 2 plots showing the actual, mean model and GRNN optimal aspect score for all test input assessment.

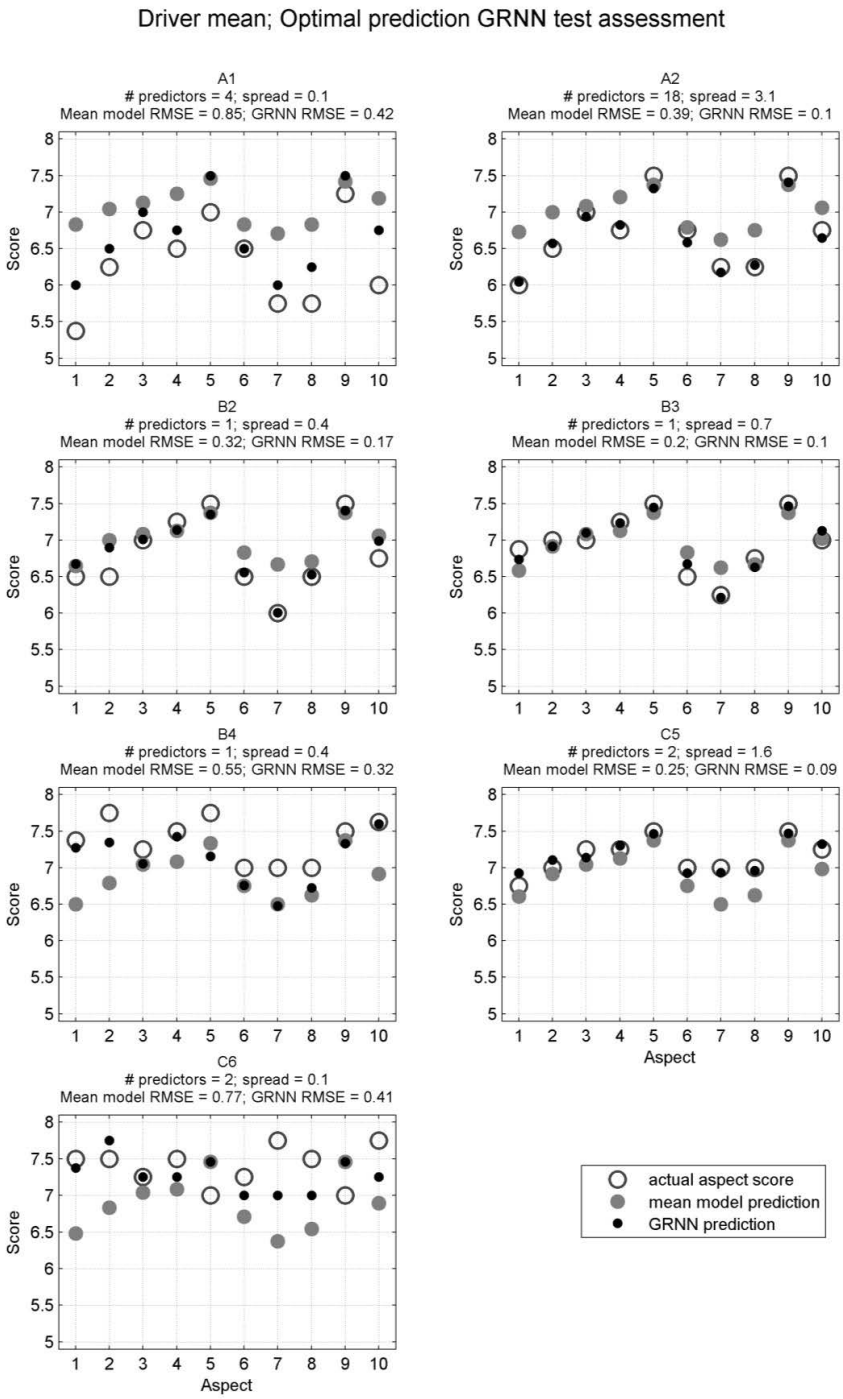


Figure 179. Driver mean plots showing the actual, mean model and GRNN optimal aspect score for all test input assessment.

A key question in the development of a tire is "How can this tire improve vehicle handling?" Good handling tires contribute not only to active safety of vehicles, but also to the pleasure of driving. Handling performance is largely determined by the driver. Therefore, the final and most important handling assessment of tires is done by professional test drivers driving on a handling circuit and giving their subjective opinion. This provides the tire manufacturer with the important tire handling performance, but it gives limited information on *what* the driver perceives as good and *how* his opinion is formed; the driver is still a 'black box'. This research presents three methods that predict the driver's opinion about tire handling, based on vehicle dynamics measures. In addition, driver's mental workload measures showed to be good indicators of driver's perceived tire handling behavior, even when the performance measures do not show differences. This provides a first step in opening up the 'black box' of the driver by answering these *what* and *how* questions.

ISBN 978-952-60-6547-2 (printed)
ISBN 978-952-60-6548-9 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University
School of Engineering
Department of Engineering Design and Production
www.aalto.fi

BUSINESS + ECONOMY

ART + DESIGN + ARCHITECTURE

SCIENCE + TECHNOLOGY

CROSSOVER

DOCTORAL DISSERTATIONS

Aalto-DD 202/2015